Bundesministerium für Bildung und Forschung

Finanziert von der Europäischen Union
NextGenerationEU

technopolis group

Fraunhofer ISI

LAW & INNOVATION

Mai 2024

Wissenschaftliche Begleitung und Vernetzung der Projekte zur Entwicklung und praktischen Erprobung von Datentreuhandmodellen in den Bereichen Forschung und Wirtschaft

Bericht zu Arbeitspaket 1.2: Anforderungen und Umsetzungshemmnisse für Datentreuhandmodelle

Mai 2024

Wissenschaftliche Begleitung und Vernetzung der Projekte zur Entwicklung und praktischen Erprobung von Datentreuhandmodellen in den Bereichen Forschung und Wirtschaft

Bericht zu Arbeitspaket 1.2: Anforderungen und Umsetzungshemmnisse für Datentreuhandmodelle

Technopolis Group: Stephan Kreutzer, Prof. Dr. Thomas Heimer, Heike Nachtigall, Lisa Pschorn, Fiona Bauer

Fraunhofer ISI: Prof. Dr. Knut Blind, Dr. Nicholas Martin

Law & Innovation: Prof. Dr. Max von Grafenstein

GRI GmbH, RWTH Aachen: Prof. Dr. Rita Streblow, Junsong Du, Joel Schölzel

Die Studie wird im Auftrag des Bundesministeriums für Bildung und Forschung (kofinanziert durch das Programm „NextGenerationEU“ der Europäischen Union) durchgeführt.

RWTH Aachen University
E.ON Energieforschungszentrum
Lehrstuhl für Gebäude- und Raumklimatechnik
Mathieustr. 10
D-52074 Aachen

Verfügbar über das Institutionelle Repositorium der RWTH Aachen University
DOI: 10.18154/RWTH-2024-04375

Inhaltsverzeichnis

Tabellen

Abbildungen

1 Abkürzungsverzeichnis

Tabelle 1 Abkürzungsverzeichnis

Abkürzung	Definition
AES	Advanced Encryption Standard
API	Application Programming Interface
BF	Begleitforschung
DGA	Data Governance Act
DT	Datentreuhänder
DTM	Datentreuhandmodell
FDZ	Forschungsdatenzentren
MVP	Minimum Viable Product
PIMS	Personal Information Management Systems
QT	Querschnittsthema

2 Glossar

Tabelle 2 Glossar

Begriff	Definition
Datensouveränität	Datensouveränität bezieht sich auf das Konzept, wonach natürliche Personen das Recht und die Kontrolle über ihre eigenen Daten haben sollten. Insbesondere sollten natürliche Personen selbst entscheiden, wie die Daten, die ihr gehören oder zu ihrer Identifizierung verwendet werden können, erhoben, gespeichert, verarbeitet und geteilt werden sollen. Die DSGVO enthält keine eindeutige Definition der Datensouveränität, aber die entsprechende Vorgabe wird bereits vorgesehen. Es wird in Art. 5 DSGVO klar darauf hingewiesen, dass zum Schutz der Rechte und Freiheiten der betroffenen Person die Verarbeitung, Erhebung und Speicherung der personenbezogenen Daten kontrolliert werden müssen.
Data Commons	Anwendung von Grundsätzen und Erkenntnissen aus der Bewirtschaftung gemeinsamer Ressourcen auf Daten (ODI, 2019). Die Anforderungen an einen DT sind häufig vergleichbar mit denen an den Verwalter eines Data Commons.
Data Space	Dezentralisiertes Datenökosystem, das auf gemeinsam vereinbarten Technologien, Werten, Standards oder Schnittstellen basiert und einen effektiven und vertrauenswürdigen Austausch von Daten zwischen den Teilnehmenden ermöglicht. Ziel ist die Schaffung mehrseitiger Märkte, die Steigerung bzw. Verbesserung der Datenverfügbarkeit und bessere Datenzugriffsmöglichkeiten der Teilnehmenden. Bausteine eines Datenraums umfassen Bedingungen und Mechanismen in Verbindung mit Datenangeboten, wie für die Preisgestaltung und Vertragsprozesse sowie die Veröffentlichung und das Auffinden von Datenangeboten (Otto et al., 2022). Innerhalb des Ökosystems werden verschiedene Rollen definiert, die sich in die drei Gruppen (i) Governance, (ii) Teilnehmende, und (iii) unterstützende Dienstleistungen einteilen lassen. In einem Datenraum, wie oben definiert, übernimmt die Governance Ebene (bei Gaia-X auch unter der Bezeichnung des

Begriff	Definition
	Föderators) Funktionen wie z.B. das Identitätsmanagement, Zertifizierung und die Orchestrierung des Ökosystems/Datenraums im Allgemeinen (Otto et al., 2022; Kraemer et al., 2023) und erinnert somit stark an unser breit angesetztes Begriffsverständnis eines Datentreuhänders (DT). Datenräume können offen oder geschlossen sein (in verschiedenen Abstufungen), der Zugang kann diskriminierend oder nichtdiskriminierend erfolgen.
Data Trust	Eine rechtliche Struktur, welche die Verantwortung und Verpflichtungen für die Verwaltung von Daten für einen vereinbarten Zweck gegenüber einer Gruppe von Begünstigten wahrnimmt (ODI, 2019, IPPI, 2022).
Data Trustee	Intermediär, der die Funktionsweise eines Data Trusts gewährleistet (s. Eintrag Data Trust). Die Anforderungen an einen Data Trustee sind häufig vergleichbar mit denen an einen Datentreuhänder.
Datenintermediär	In der allgemeinen Diskussion ist der Begriff des Datenintermediärs ein Sammelbegriff für Modelle oder Instrumente, die darauf abzielen, den Datenaustausch zwischen (mindestens zwei) Akteuren zu vereinfachen, durchzuführen oder zu begleiten. Unter die Leistungen eines Datenintermediärs fallen z.B. das Gewähren eines sicheren Datenaustauschs oder das Bekanntmachen verfügbarer Daten (DIHK, 2023b). Der Begriff des Datentreuhänders ist eine Sonderform bzw. eine Unterkategorie von Datenintermediären und zeichnet sich dadurch aus, dass er das Interesse aller beteiligten Akteure wahrt. Dies umfasst z.B. die Gewährleistung des Datenschutzes, des Geschäftsgeheimnisschutzes oder der Datennutzbarkeit (DIHK, 2023b). Im Sinne von Art. 2 Nr. 11 DGA ist ein Datenintermediär ein Dienst zur Herstellung einer Geschäftsbeziehung zwischen einer unbestimmten Anzahl von Datengebenden und Datennutzenden zur Ermöglichung gemeinsamer Datennutzung. Eine Diskussion der rechtlichen Perspektive findet sich in Abschnitt 6.
Datenökosystem	Ein Datenökosystem wird im Folgenden als ein Umfeld begriffen, in dem verschiedene Akteure zusammenkommen, um Daten zu produzieren, anzubieten, zu finden und zu „konsumieren“ (d.h. nachzunutzen, zu verarbeiten, anzureichern, zu archivieren, zu publizieren, Entscheidungen darauf zu fällen, etc.) (Putnings, 2021).
Datengenossenschaft (data cooperative)	Datengenossenschaften sind Organisationen, die Daten von ihren Mitgliedern sammeln, die zusammengefassten Daten verarbeiten und monetarisieren und die Mitglieder für ihre individuellen Beiträge entschädigen. Diese Genossenschaften schaffen ein Ökosystem des Vertrauens zwischen ihren Mitgliedern (Mehta et al., 2021). Die Mitglieder der Datengenossenschaft üben demokratisch die Kontrolle und Entscheidungsfindung über ihre Daten aus (ODI, 2019).
Domäne	z.B. Mobilität, Energie, Klima
Datentreuhänder	Neutraler Intermediär, der einen vertrauensvollen und fairen Ausgleich der Interessen der beteiligten Akteurinnen und Akteure – Datengebende sowie Datennutzende – ermöglicht, gegebenenfalls neue Vertrauensbeziehungen anbahnt, den technischen und organisatorischen Zugang zu qualitativ hochwertigen Daten unter Wahrung des Datenschutzes sowie Interoperabilität garantiert (BMBF, 2021).
Datentreuhandmodell	Spezifische technische, rechtliche und organisatorische bzw. geschäftliche Ausformung eines Datentreuhänders.
Neutralität unter dem DGA	Neutralitätsverpflichtung für Datenvermittlungsdienste (siehe unten) aus Art. 12 lit. a) DGA, d.h. klare und strukturelle Trennung von Datenvermittlung und -nutzung (Erw.gr. 32). Der Datenvermittlungsdienst muss zur Vorbeugung von Interessenkonflikten von einer unabhängigen jur. Person erbracht werden, d.h. rechtlich unabhängig sein. Wirtschaftliche Unabhängigkeit wird dagegen nicht verlangt. Der Anbieter des Datenvermittlungsdiensts darf jedoch seine Leistung nicht von der Nutzung anderer eigener Dienste abhängig machen vgl. Art. 12 lit. b) DGA und Erw.gr.33 – Kopplungsverbot (Hennemann/von Ditfurth, 2022). Demgegenüber verlangt § 26 Abs. 1 Nr. 2 TTDSG für den Bereich der Einwilligungsagenten („Anerkannte Dienste zur Einwilligungsverwaltung“), dass der Einwilligungsagent „kein wirtschaftliches Eigeninteresse an der Erteilung der Einwilligung und

Begriff	Definition
	an den verwalteten Daten haben (darf) und unabhängig von Unternehmen (… / sein muss), die ein solches Interesse haben können". Ob damit ein wirtschaftliches Eigeninteresse noch möglich ist, ist zweifelhaft.
Datenmodell	Ein Datenmodell ist die formale Abbildung der Informationsobjekte der betrachteten Diskurswelt mittels ihrer Attribute und Beziehungen. Ziel ist die eindeutige Definition und Spezifikation der in einem Informationssystem zu verwaltenden Objekte, ihrer für die Informationszwecke erforderlichen Attribute und der Zusammenhänge zwischen verschiedenen Informationsobjekten (Gabler, 2023).
Personal Information Management Systems	Softwarelösungen, die es Einzelpersonen ermöglichen, ihre persönlichen Daten sicher zu verwalten und sie nach eigenem Ermessen mit anderen zu teilen. Anbieter von Online-Diensten und Werbetreibende müssen mit PIMS interagieren, wenn sie diese Daten verarbeiten wollen. Bei PIMS handelt es sich um Geschäftsmodelle, die im Auftrag von Verbraucherinnen und Verbrauchern gegenüber Dritten agieren (Schneider, 2022). PIMS wird im Rahmen der Begleitforschung als möglicher technischer Baustein eines Datentreuhänders verstanden, aber nicht als ein Datentreuhändermodell.
Querschnittsthema	Querschnittsthemen in der Begleitforschung sind (i) technische Infrastruktur, (ii) rechtliche Rahmenbedingungen, (iii) Geschäftsmodelle, (iv) Akzeptanz, Skalierung und Transfer.
Sektor	Wirtschaft, Wissenschaft, Zivilgesellschaft, Politik und Verwaltung.
Transaktionsbasierter Datentreuhänder	Datentreuhänder, bei dem keine dauerhafte Datenspeicherung erfolgt. Die Datenverarbeitung findet transaktionsbasiert bzw. anlassbezogen statt. Ein Beispiel ist das Projekt EuroDat. Das Gegenteil, ein Datentreuhänder, bei dem eine dauerhafte Datenspeicherung erfolgt, wird in der Literatur auch als „Silo-Datentreuhänder" bezeichnet (Reiberg et al., 2023).

3 Executive Summary

Der vorliegende Bericht präsentiert die Ergebnisse aus Arbeitspaket (AP) 1.2 – Anforderungen und Umsetzungshemmnisse in Datentreuhandmodellen (DTM) in der Begleitforschung (BF) zu den vom Bundesministerium für Bildung und Forschung (BMBF) geförderten „Projekten zur Entwicklung und praktischen Erprobung von Datentreuhandmodellen in den Bereichen Forschung und Wirtschaft“. Die Befunde bauen auf der Auswertung wissenschaftlicher Literatur im Bericht zu AP 1.1 auf und beruhen auf Interviews mit Projektleitenden (zum Teil im Beisein weiterer Projektmitarbeitender) einer Online-Befragung und Workshops mit Vertreterinnen und Vertretern der geförderten Projekte, sowie auf Interviews mit externen Expertinnen und Experten.

Zentrale Erkenntnis ist, dass DTM ein Instrument darstellen, welches Anreize für das Datenteilen innerhalb und zwischen Wissenschaft und Wirtschaft schaffen und die verschiedenen Interessen von Datennutzenden und -gebenden ausgleichen kann. Dabei zeichnet sich ab, dass es kein universelles DTM geben wird, sondern dass sich vielmehr ein Werkzeugkasten an Services und Governance-Modellen etabliert, der auf verschiedene Domänen übertragbare Lösungen in der Praxis erlaubt.

Eine zentrale Voraussetzung für Datentreuhandservices ist eine solide technische Infrastruktur. DTM verwenden eine Vielzahl von technischen Architekturen und Bausteinen. Die Festlegung von generischen und wiederverwendbaren technischen Bausteinen wäre an dieser Stelle wünschenswert, gestaltet sich durch branchenspezifische Anforderungen und rechtliche Beschränkungen jedoch schwierig. Die Interoperabilität von DTM erfordert einheitliche Schnittstellen, Datenformate und Ontologien. Für die Datensicherheit und -souveränität sind Identitäts- und Zugriffsmanagement, Anonymisierung, Pseudonymisierung und Datenverschlüsselung von großer Bedeutung.

Fehlende Rechtssicherheit in Bezug auf Datenschutz und Geschäftsgeheimnisse wird in vielen Förderprojekten als Umsetzungshemmnis wahrgenommen. Schwierigkeiten bereiten hier insbesondere die Frage zur technisch-organisatorischen Umsetzung rechtlicher Anforderungen und Haftungsfragen. Eine der wichtigsten Ansätze, den extremen Aufwand in Sachen Koordination und technisch-organisatorischer Implementierung und so die Kosten zu reduzieren, ist die Standardisierung der Lösungsmodelle. Eine Systematisierung der verschiedenen DTM-Ansätze wird einer der Schwerpunkte der weiteren BF sein.

Geschäftsmodelle für DT müssen nicht nur finanziell tragfähig sein, sondern auch Neutralitätsanforderungen konkret adressieren. Die durch die Regulierung vorgegebene Neutralitätsanforderungen stellen eine Herausforderung für das Geschäftsmodell der privatwirtschaftlich finanzierten DT dar, da sie die Wertschöpfungsmöglichkeiten einschränken. Auch ist bisher nicht juristisch eindeutig, wie weit die Neutralitätsanforderungen reichen und welcher Spielraum bei den DT in der Entwicklung von Geschäftsmodellen haben.

Die Steigerung der Akzeptanz von DTM ist von entscheidender Bedeutung für deren Erfolg. Wichtigste Treiber sind hierbei Vertrauen in die Sicherheit des DT und den Umgang mit den geteilten Daten, sowie der Nutzen, den Datengebende und -nutzende aus der Verarbeitung der Daten ziehen. Überzeugende exemplarische Use-Cases aber auch Standards und Zertifizierungen können dabei helfen, Vertrauen bei den Datennutzenden und -gebenden zu schaffen. Skalierung wird von den Förderprojekten eher graduell angegangen.

Im nächsten Schritt wird die BF die in den Pilotprojekten erprobten Lösungsansätze verstärkt in den Blick nehmen.

4 Einführung: Einordnung in die Begleitforschung und methodisches Vorgehen

Der **vorliegende Bericht** präsentiert die Ergebnisse aus dem AP 1.2 und untersucht die **Anforderungen** an die in den Förderprojekten zu entwickelnden DTM sowie mögliche **Hemmnisse** bei ihrer Umsetzung.

Ausgangspunkt ist die **Arbeitshypothese**, dass Daten aus gesamtgesellschaftlicher und insbesondere innovationspolitischer Sicht aufgrund verschiedener Interessenlagen und Informationsasymmetrien zwischen Datengebenden und Datennutzenden nur unzureichend innerhalb und zwischen Anwendungsdomänen geteilt werden. Die Begleitforschung geht hierbei davon aus, dass dies dem Umstand geschuldet ist, dass viele Dateninhabende die mit dem Datenteilen verbundenen Risiken stärker gewichten als den potenziellen Nutzenden (**Wert-Risiko-Dilemma**). Nach Auffassung der Begleitforschung können **DTM als ein Instrument zur Förderung des Teilens von Daten zwischen und innerhalb von Anwendungsdomänen** dienen. Hieraus leitet sich die zentrale **Leitfrage** ab, welche Herausforderungen sich DT in der Praxis stellen und wie ein DT ausgestaltet sein muss, um das Datenteilen tatsächlich zu fördern. Die Herangehensweise an diese zentrale Fragestellung wird anhand von **vier Querschnittsthemen** (QT) untergliedert: QT1 Technische Infrastruktur, QT2 Rechtliche Rahmenbedingungen, QT3 Geschäftsmodelle, QT4 Akzeptanz, Skalierung und Transfer.

Im Rahmen einer Literaturanalyse wurden in AP1.1 anhand der vier QT bereits bestehende Erkenntnisse zusammengetragen und analysiert. Tabelle 3 fasst die in der Literaturanalyse identifizierten zentralen Herausforderungen für DT zusammen.

Tabelle 3 Zentrale Herausforderungen von DT aus der Literatur

	Forschungsstand	Identifizierte Herausforderungen	Forschungslücken bzw. Forschungsbedarf
QT1	Einheitliche Grundfunktionalitäten wurden definiert.	• Sich entwickelnde Marktanwendungen, die technische Anforderungen weiter verändern • transparente und kontrollierbare Verteilung von Rechten und Interessen	• Datensicherheit, -kontrolle und -souveränität und zugehörige Servicelevel • Technische Maßnahmen 1) zum Schutz sensibler Daten, 2) zur Kontrolle der Datennutzung und 3) zur Gewährleistung der Datensouveränität
QT2	Umfassende Diskussion in der Literatur zur rechtlich-organisatorischen Struktur von DTM und ihrer Abgrenzung sowie ansatzweise zu den Funktionen von DTM und den zu lösenden Problemen.	• Hohe Rechtsunsicherheit und fehlende Skalierbarkeit bei der Rechtsanwendung führen unter Akteuren zum sogenannte Wert-Risiko-Dilemma (d.h. zur Wahrnehmung, dass die erwartete Wertschöpfung beim Datenteilen niedriger ist als die Compliance-Risiken und die damit verbundenen Kosten)	• Funktionen, die DT erfüllen sollen, insbesondere, wie DTM-Akteure dabei unterstützen können, die Rechtsunsicherheit auf kostensparsame Weise zu reduzieren, so dass das Datenteilen wieder lohnenswert erscheint
QT3	Breite Literatur zu möglichen Geschäftsmodellkomponenten und Funktionen sowie zur Problematisierung der Neutralitätsanforderungen	• Eingeschränkter Spielraum für Geschäftsmodellansätze durch Neutralitätsanforderungen; • Etablierung nachhaltiger Finanzierungskonzepte von DT; • Unklarheiten über die Ausgestaltung und die	• Geschäftsmodellausgestaltung unter Neutralitätsbedingungen: • (1) Nachfrageseite /Nutzendeseite: • Nachfrage nach DTM im Ökosystem, Rolle von DTM im Ökosystem (z.B. Friktionen, Risiken, Potenziale);

	Forschungsstand	Identifizierte Herausforderungen	Forschungslücken bzw. Forschungsbedarf
		Erfolgsfaktoren von DT-Geschäftsmodellen im Allgemeinen durch das Fehlen bereits etablierter Geschäftsmodelle und Konzepte in der Praxis	• Anreize für Datennutzende (potenzielle Nutzende von DTM); • monetäre und nicht monetäre Anreize für Datengebende; • (2) Angebotsseite: • Anreize für potenzielle Betreiber von DTM; • Finanzierungskonzepte von DT, inkl. Bepreisung von Daten oder Dienstleistungen; • Geschäftsmodellkomponenten bzw. Funktionen (inkl. Bepreisung/Finanzierung)
QT4	• begrenzte Forschung zu Akzeptanz und Skalierung bei DT (v.a. graue Literatur); • vergleichsweise mehr Forschung zu Akzeptanz von Datenteilen allgemein; • kaum zu Standards, Zertifizierung, FRAND-Bedingungen: Fair, Reasonable and Non Discriminatory	• Prozesse, um Akzeptanz proaktiv zu schaffen bzw. Dilemma kollektiven Handelns aufzulösen; • erst wenige Standards & keine Zertifizierungen, kurzfristiger Bedarf allerdings unklar	• was die konkreten Hemmnisse und Treiber für Akzeptanzaufbau sind; • welche Prozesse verwendet werden, um Akzeptanz zu schaffen und wie diese skalieren können; • inwiefern der Staat zum Akzeptanzaufbau beitragen kann; • tatsächlicher aktueller Bedarf nach Standards bzw. Zertifizierungen; • Anreize, Standards bzw. Zertifizierungen entwickeln; • welche Stellen aus Sicht der Akteure als Zertifizierer in Frage kommen; • möglicher Beitrag von staatlichen / staatlich unterstützten Infrastrukturen für Akzeptanz und Skalierung

Der Fokus der BF in AP1.2 lag darauf, zum einen neue Erkenntnisse bezüglich des bereits identifizierten Forschungsbedarfs zu gewinnen, insbesondere zu projektübergreifenden, zentralen Herausforderungen und Umsetzungshemmnissen sowie zur Ausgestaltung von DTM, und zum anderen einen Ausblick auf innovative Lösungsansätzen zu geben. Darüber hinaus wurde der aktuelle Arbeitsstand der Förderprojekte ermittelt.

Folgende Daten wurden für diesen Bericht erhoben:

1. **Leitfadengestützte, explorative Interviews mit externen Expertinnen und Experten** (drei je QT, insgesamt zwölf) zum Abgleich der Erkenntnisse aus der Literaturauswertung in der Bestandsaufnahme (AP 1.1) und Identifikation weiterer wichtiger Aspekte und Fragestellungen. Die so gewonnenen Erkenntnisse dienten der Feinkonzeption der Interview-Leitfäden und Online-Fragebögen und flossen direkt in die folgenden Kapitel ein.
2. Eine Auswertung der Gesamtvorhabenbeschreibungen und Zwischenberichte der Förderprojekte sowie die Aufbereitung zentraler Erkenntnisse in Form von **Factsheets**. Diese dienten auch der Vorbereitung der weiteren Befragungen der Förderprojekte.
3. **Leitfadengestützte Interviews mit den Gesamtverantwortlichen und Fachleuten aus den Förderprojekten** zur Ermittlung (i) des aktuellen Arbeitsstands der Projekte, (ii)

der zentralen Herausforderungen und Lösungsansätze in Bezug auf Fragestellungen entlang der vier QT, letzteres je nach individueller Schwerpunktsetzung der Projekte. Die Erkenntnisse aus diesen Interviews flossen direkt in die folgenden Kapitel ein und dienten zur Konzeption der Onlinebefragung. Die Auswertung der Interviews erfolgte anonymisiert.

4. **Eine Online-Befragung von Fachleuten in den Förderprojekten** zur Validierung der aus den Interviews gewonnen Erkenntnisse und zur Ermittlung eines möglichst repräsentativen Meinungsbilds innerhalb der Projekte. Die Online-Befragung bestand aus insgesamt fünf Modulen mit größtenteils geschlossenen Fragen: einem allgemeinen Einführungsmodul sowie je einem für jedes QT. An der Umfrage beteiligten sich 19 der 20 Projekte, wobei diese nicht alle Fragen beantworteten. Da nicht alle Projekte zu allen Survey-Fragen antworteten, wird in den folgenden Kapiteln beim Verweis auf Aussagen zu einzelnen Survey-Fragen auch die Anzahl der sich beteiligten Projekte angegeben.
5. **Drei Fachgruppenworkshops** (einer mit Schwerpunkt auf QT 1, einer mit Schwerpunkt auf QT 2 und einer mit Schwerpunkt auf QT 3 und 4 gebündelt) dienten der Präsentation und gemeinsamen Diskussion vorläufiger Befunde aus der BF mit Vertreterinnen und Vertretern der Förderprojekte und der Identifikation weiterführender Forschungsfragestellungen.

Im Anschluss der Befragungen (Interviews, Onlinebefragungen) sowie der Fachgruppenworkshops fand ein Workshop (online) im Kreis des Teams der BF sowie des Projektbeirats statt, in dem die Ergebnisse von AP1.2 diskutiert wurden. Diese Erkenntnisse flossen ebenfalls in den vorliegenden Bericht ein.

Im Folgenden werden die Befunde zunächst für jedes QT präsentiert. Anschließend werden übergeordnete Schlussfolgerungen zusammengefasst und ein Ausblick auf die nächsten Schritte in der BF gegeben.

5 QT1: Technische Infrastruktur und Datensicherheit

Die inhaltliche Schwerpunktsetzung in QT 1 liegt in AP 1.2 auf den Anforderungen und Herausforderungen bei der technischen Infrastruktur und der Datensicherheit. Insbesondere werden die in AP1.1 identifizierten Forschungslücken adressiert, die in der übergreifenden Einführung in Kapitel 1, insbesondere in Tabelle 1, dargestellt wurden. Die inhaltliche Schwerpunktsetzung lässt sich in drei übergreifende Themenblöcke unterteilen:

Ausgestaltung der technischen Infrastruktur inkl. der einzelnen Bausteine und der Architektur des DTM,

Gestaltung von Schnittstellen und Instrumenten zur Herstellung von Datenportabilität und Interoperabilität verschiedener Dienste,

Entwicklung von Instrumenten zur Gewährleistung einer effektiven und transparenten **Verwaltung von Datenzugriffsrechten** und der **Kontrolle ihrer Einhaltung (Usage Control)** und der **Datensicherheit**.

5.1 Aktueller Forschungsstand der Förderprojekte

Eine sorgfältige Analyse der **Vorhabenbeschreibungen** sowie der **Zwischenberichte** sämtlicher Pilotprojekte haben Einblicke in den jeweiligen aktuellen Forschungsstand ergeben.

Sämtliche Pilotprojekte greifen in unterschiedlichem Ausmaß auf vorangegangene Forschungsarbeiten und Erfahrungen, bereits existierende Tools und Services oder auf Open-Source-Ansätze zurück. Lediglich knapp **30 % der Pilotprojekte** betreiben **Grundlagenforschung**, wie beispielsweise die Erforschung mathematischer Methoden oder die Konzeptionierung neuartiger Verschlüsselungsansätze, oder widmen sich der **Entwicklung von neuen technischen Bausteinen**. Davon konzentrieren sich zwei Pilotprojekte auf Verschlüsselungsverfahren, die in Kapitel 5.2 noch näher erläutert werden. Drei Projekte beschäftigen sich mit der Klassifizierung der sensiblen Daten in verschiedene Sicherheitsklassen je nach Risiko, sodass angemessene Sicherheitsmaßnahmen implementiert werden können. Zudem zielen zwei Pilotprojekte darauf ab, ein Bewertungsverfahren der Nutzendereputation zu entwickeln, um eine dynamische Datenfreigabe zu ermöglichen. Ein weiteres Pilotprojekt sieht die Entwicklung einer Ontologie für die (Wind-)Energiebranche vor, was Datenportabilität und -interoperabilität durch die Beschreibung von Metadaten in Form einer Ontologie ermöglichen soll.

Die Mehrheit der Pilotprojekte legt ihren Fokus auf die Implementierung einzelner DTM-Module auf Grundlage bereits existierender Ansätze sowie auf deren Kombination zu einer umfassenden Treuhandstelle. Da für einzelne Funktionalitäten (z.B. Autorisierung, Authentifizierung oder Identitätsmanagement) bereits weitgehend ausgereifte Ansätze existieren, wird die Frage, wie ein reibungsloses Zusammenspiel dieser Komponenten gewährleistet werden können, im Mittelpunkt der aktuellen Forschung stehen. Dementsprechend greifen mindestens 25 % der Pilotprojekte auf die ***International Dataspaces* (IDS)-Referenzarchitektur** zurück, insbesondere auf die IDS-Konnektoren für die Datenbereitstellung und Datennutzungskontrolle. Diese Architektur eines Dataspace wird anhand von dem Ansatz eines Pilotprojektes in Kapitel 5.2 genauer erläutert. Darüber hinaus sind **die Entwicklungen der GAIA-X Initiative** für die meisten Pilotprojekte von Interesse. Insgesamt haben elf Pilotprojekte GAIA-X in ihren Vorhabenbeschreibungen erwähnt. Die anderen Pilotprojekte sind GAIA-X gegenüber aber auch prinzipiell offen. Eines der Projekte hat deutlich gemacht, dass sie den Fortschritt des GAIA-X-Projekts beobachten und gegebenenfalls die Anforderungen ihres Projektes entsprechend anpassen. Eine Herausforderung für alle Pilotprojekte könnte die Notwendigkeit darstellen, den

Entwicklungsstand bestehender technischer Bausteine, offener Standards und Initiativen ständig zu überprüfen.

In der Literaturanalyse für AP1.1 wurde die Art der **Datenspeicherung (zentral oder dezentral)** als eine der wichtigsten Faktoren für DTM identifiziert, denn die technische Implementierung der meisten Funktionalitäten kann maßgeblich dadurch beeinflusst werden. Die meisten der untersuchten Studien rekurrieren auf zentrale Ansätze. Die dezentralen Ansätze werden erst in kürzlich veröffentlichten Studien untersucht. Diese Tendenz spiegelt sich gut in den Pilotprojekten wider. Circa die Hälfte der Pilotprojekte verfolgt dezentrale oder föderierte Ansätze. Bei einem Pilotprojekt stellt die Erforschung der Dezentralität ein Hauptziel dar, weil in der adressierten Branche bisher nur zentrale Technologien zum Einsatz kommen. Einige Pilotprojekte setzen sich sogar mit dem Konzept „*algorithm to data*" auseinander. Dieses Konzept läuft auf eine noch radikalere Dezentralisierung hinaus, da nicht nur die Daten dezentral gespeichert werden, sondern auch die Datenverarbeitung dezentralisiert wird.

Viele der Pilotprojekte bearbeiten ähnliche Forschungsfragen. Um die Nutzbarkeit der Daten zu gewährleisten, setzen zehn Projekte den Schwerpunkt auf die Harmonisierung heterogener Daten, während weitere acht Projekte Maßnahmen zur Integration externer Partner und Datenbanken durch die Implementierung von Datenschnittstellen ergreifen. Zusätzlich legen acht Pilotprojekte besonderen Wert auf die Entwicklung von Methoden zur Bewertung und Gewährleistung der Datenqualität. Im Kontext der Datensicherheit und -souveränität konzentrieren sich wiederum acht Projekte auf den Aufbau von Zugriffsmanagementsystemen, während weitere acht Projekte technische Ansätze im Bereich des Einwilligungsmanagements verfolgen. Des Weiteren widmen sich sieben Pilotprojekte der Pseudonymisierung und Anonymisierung von Daten, wobei der Schwerpunkt auf der Entwicklung automatisierter Verfahren für komplexe Datensätze lag. Es wurde deutlich, dass bei sämtlichen individuellen Forschungsfragen und technischen Umsetzungen beträchtliche Synergien zwischen den Pilotprojekten geschaffen werden können. Dies unterstreicht die Relevanz der Zusammenarbeit und des Wissensaustauschs (z.B. der Fachgruppeworkshop in Rahmen von AP 2) innerhalb des Forschungsverbunds, um die Herausforderungen effektiv anzugehen.

Abschließend erfolgte eine detaillierte Analyse des gegenwärtigen Fortschritts der Pilotprojekte im Kontext von QT1. Auf Grundlage der betriebenen Literaturrecherche haben die meisten Projekte bereits mit der **Konzeptionierung einzelner Module sowie der Gesamtarchitektur begonnen**. Dies beinhaltete unter anderem die Festlegung von Standards und Schnittstellen mit den Projektpartnern. Zudem wurde beispielweise die **Harmonisierung ausgewählter Datensätze** initiiert, wodurch diese nun für Testzwecke zur Verfügung stehen. Die weitere Vorgehensweise ist in den meisten Pilotprojekten klar strukturiert und fokussiert sich auf die konkrete Umsetzung der erarbeiteten Konzepte. Jedoch ist hervorzuheben, dass **einige Pilotprojekte bereits konkrete Prototypen und Demonstratoren entwickelt haben, während sich andere noch in der Planungsphase befinden.**

Zusammenfassend lässt sich feststellen, dass fast alle Pilotprojekte auf bestehende Arbeiten und Ansätze zurückgreifen, wobei lediglich etwa ein Drittel grundlegende Forschung betreiben und neue technische Elemente entwickeln. Ihre **Forschungsschwerpunkte** umfassen Verschlüsselungsmethoden, Sicherheitsklassifizierung und die Bewertungsmethoden der Datenqualität und Reputation. Die Hauptbemühungen der Pilotprojekte konzentrieren sich auf die Integration vorhandener technischer Bausteine, wobei die IDS-Referenzarchitektur in vielen Pilotprojekten genutzt wird und die Entwicklungen von GAIA-X aufmerksam verfolgt werden. Hervorzuheben sind insbesondere die Forschungsrichtungen zur Sicherung der Datenverfügbarkeit sowie zur Gewährleistung von Datensicherheit und -souveränität. Die Anforderungsanalyse und Literaturrecherche wurden weitestgehend von den Pilotprojekten

abgeschlossen und die praktische Umsetzung und Implementierung von Prototypen hat teilweise begonnen.

5.2 Ausgestaltung der technischen Infrastrukturen

Die Literaturanalyse von AP1.1 zeigte, dass es eine große Anzahl von Konzepten gibt, die mit DTM in Verbindung stehen, und dass es meist keine klare Abgrenzung zwischen diesen gibt. Zu den gebräuchlichsten Konzepten gehören Data Space und Datenintermediär. Diese nicht klar abgegrenzten Konzepte stellen zum einen eine Herausforderung für den technischen bzw. akademischen Austausch dar, da die unterschiedliche Verwendung von Begrifflichkeiten zu Missverständnissen führen kann. Zum anderen erhöht diese Unklarheit auch die Eintrittsbarrieren in die DTM-Branche, da die verschiedenen sich überschneidenden, aber unterschiedlichen technologischen Ansätze die Lernkurve im Vorfeld abflachen.

Die BF hat diese Ausgangssituation aufgegriffen, und die verschiedenen technischen Infrastrukturen, die sich in den Förderprojekten herausgebildet haben, untersucht, um ihre Gemeinsamkeiten oder Einzigartigkeiten herauszufinden. Zu diesem Zweck wird sich Kapitel 5.2.1 auf die verschiedenen Architekturen des DTM konzentrieren. Im Anschluss wird in Kapitel 5.2.2 auf die Übertragbarkeit und Wiederverwendbarkeit der technischen Bausteine eingegangen.

5.2.1 Die Architektur des DTM

In AP1.2 wurden in den Interviews, der Online-Befragung, der Analyse der Fortschrittsberichte und der Diskussion im Fachgruppeworkshop unterschiedliche Architekturen benannt. **Die am häufigsten auftauchenden Ansätze sind Data Space und Datenintermediär.**

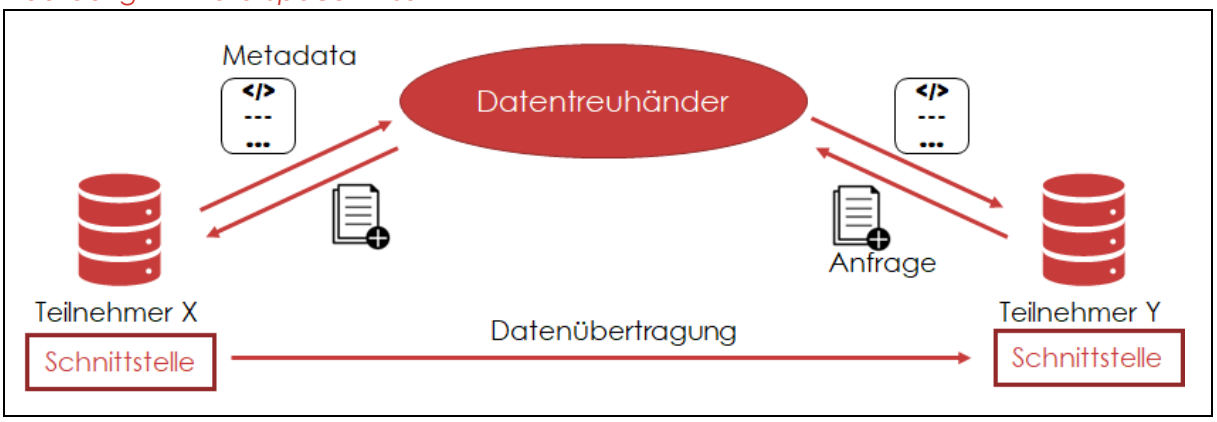
Data Space:

Ein Pilotprojekt hat seinen technischen Aufbau dargestellt, der auf dem Data Space Ansatz basiert. Wie in Abbildung 1 dargestellt, besteht der Data Space aus Teilnehmenden, die die Daten nutzen oder zur Verfügung stellen, und DT, die die Infrastruktur für den Datenaustausch bereitstellen. Diese Architektur bildet ein dezentralisiertes Datenökosystem ab, das auf gemeinsam vereinbarten Technologien, Standards oder Schnittstellen basiert.

In diesem dezentralisierten System wird jeder Teilnehmende mit der standardisierten Schnittstelle eingerichtet und kann somit eine direkte Verbindung zwischen den Datengebenden und Datennutzenden während der Datenübertragung hergestellt werden.

Die DT tragen dazu bei, ein Metadatenkatalog bereitzustellen, indem die Metadaten zentral gesammelt und geteilt werden können. Bei Interesse werden die DT auch bei der Verhandlung für einen bestimmten Datenaustausch helfen.

Abbildung 1 Data Space Ansatz



Quelle: Eigene Darstellung (GRI GmbH, RWTH Aachen)

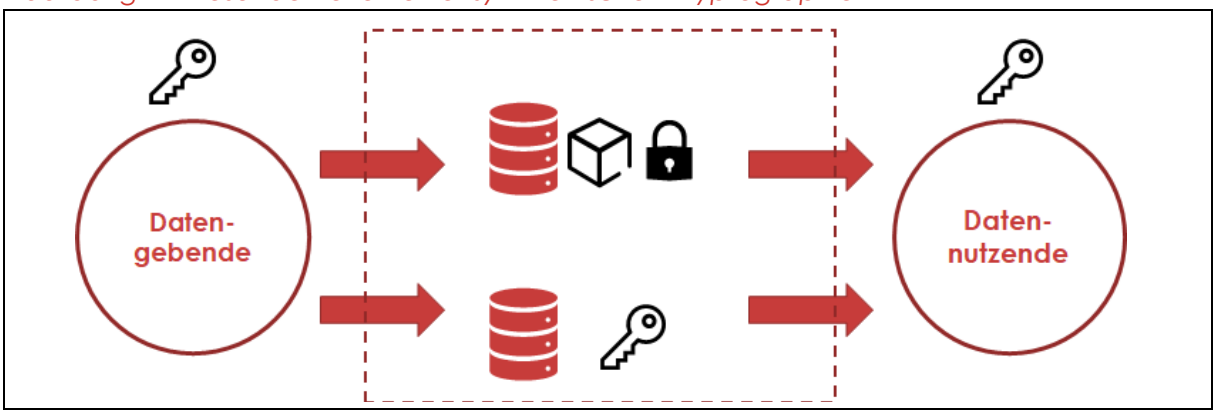
Zudem hat ein befragter Experte eine weitere Implementierungsform eines Data Space erläutert und die Herstellung der Vertrauenswürdigkeit als die Hauptverantwortung des DT betont. Die DT erstellen eine Liste der vertrauenswürdigen Teilnehmenden, in der ihre Informationen sowie technische Daten, z.B. Zertifikate oder Public Keys gespeichert und verwaltet werden. Es wurde auch vom Experten bestätigt, dass die Verwendung gemeinsamer Spezifikationen und standardisierter Schnittstellen die grundlegende Voraussetzung für ein Data Space ist. Aber die konkrete Implementierung der einzelnen technischen Bausteine wird dabei offengehalten.

Datenintermediär:

Nach Art. 2 Nr. 11 DGA ist ein Datenintermediär ein Dienst zur Herstellung einer Geschäftsbeziehung zwischen einer unbestimmten Anzahl von Dateninhabenden und Datennutzenden. Zu den Hauptverantwortungen gehören der sichere Datenaustausch und das Bekanntmachen verfügbarer Daten. Im Gegensatz zum Data Space ist die Architektur des Datenintermediärs mit zentraler Infrastruktur aufgeprägt. Der grundliegende technische Baustein eines Datenintermediäres ist ein Datenbanksystem. In der Regel werden die Daten von Datengebenden an den Datenintermediär übertragen, der sie dann an die Datennutzenden weiterleitet. Bei dieser Architektur gibt es keine direkte Verbindung zwischen den Datengebenden und den Datennutzenden, wodurch die Privatsphäre der Datengebenden geschützt wird. Allerdings werden die Daten während der Übertragung in dem Datenbanksystem des Datenintermediäres gespeichert oder zwischengespeichert, was neue Herausforderungen für die Datensicherheit und -souveränität mit sich bringt. In der Online-Befragung gaben zehn von 19 Projekten an, dass der DT die Daten auch ansehen darf, während die restlichen Projekte dies verneinten. Aus diesem Ergebnis lässt sich nicht klar schlussfolgern, ob der DT die Daten sollte ansehen dürfen. Für den Fall, dass der Zugriff des DT auf die Daten kontrolliert werden soll, haben zwei Pilotprojekte in den Interviews die technische Infrastruktur des Systems ihres Datenintermediärs vorgestellt.

Ein Pilotprojekt beabsichtigt, die symmetrische Kryptographie anzuwenden. In ihrem System werden die Daten von Datengebenden mittels symmetrischer Kryptographie verschlüsselt. Abbildung 2 zeigt die Besonderheit von ihrem System, dass die Datentreuhandstelle in zwei Parteien aufteilt. Ein DT bewahrt die verschlüsselten Daten auf, während der Schlüssel zur Entschlüsselung der Daten bei einem anderen DT gespeichert ist. Die zwei DT müssen unabhängig beauftragt werden, sodass die Daten während der Speicherung nicht verwendet werden können. Diese Auslegung hat die Herausforderung von der technischen Ebene auf die organisatorische Ebene verlagert.

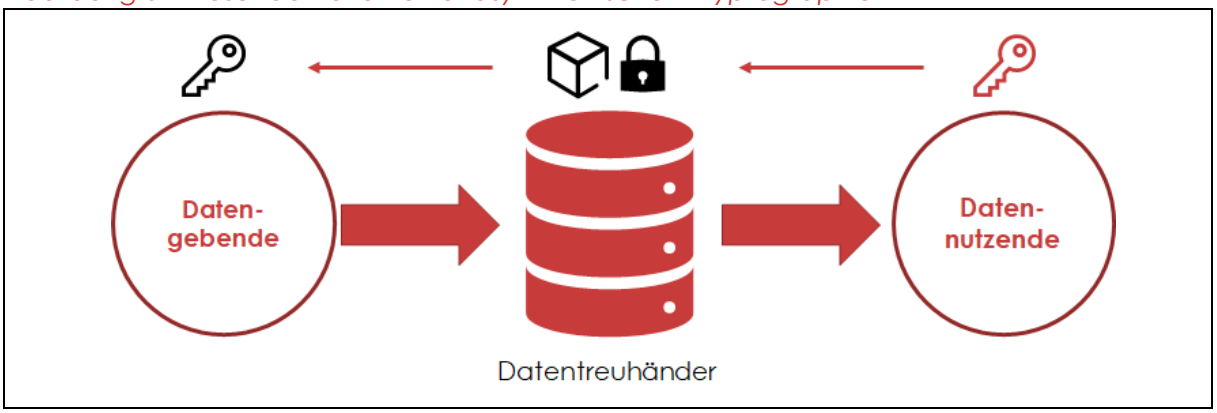
Abbildung 2 Besonderheiten einer symmetrischen Kryptographie



Quelle: Eigene Darstellung (GRI GmbH, RWTH Aachen)

Ein anderes Pilotprojekt hat zusätzlich noch die **asymmetrische Kryptographie** in seinem System integriert. Die symmetrisch verschlüsselten Daten werden auch beim DT abgelegt. Der Schlüssel bleibt allerdings erstmal auf der Seite des Datengebenden. Bei Bedarf erzeugen die Datennutzenden mit Hilfe der asymmetrischen Kryptographie ein Schlüsselpaar, in dem der private Schlüssel (*Private-Key*) lokal gespeichert und der öffentliche Schlüssel (*Public-Key*) sowie eine Datenabfrage über den DT an die passenden Datengebenden weitergeleitet werden. Nach Bewilligung der Datenabfrage werden die Datengebenden den Schlüssel für die Daten mit dem *Public-Key* verschlüsseln. Der verschlüsselte Schlüssel wird vom DT zusammen mit den verschlüsselten Daten an die Datennutzenden zurückgeschickt. Die letzteren können ihren *Private-Key* zur Entschlüsselung des Datenschlüssels verwenden, um die Daten zu entschlüsseln und zu verwenden.

Abbildung 3 Besonderheiten einer asymmetrischen Kryptographie



Quelle: Eigene Darstellung (GRI GmbH, RWTH Aachen)

5.2.2 Generische und wiederverwendbare technische Bausteine

Es lässt sich im Kapitel 5.2.1 bestätigen, dass die technische Implementierung von DTM vielfältig sein kann. Diese Vielfalt an Umsetzungsmöglichkeiten kann aber zu Schwierigkeiten bei der technischen Implementierung für einen bestimmten Anwendungsfall führen , wenn das dafür benötigte Wissen nicht systematisch zusammengefasst ist und nicht klar ist, welche Architektur am besten die Anforderungen erfüllen kann. In diesem Zusammenhang wies ein befragtes Pilotprojekt daraufhin, dass eine Datentreuhandstelle **in frühen Entwicklungsphasen noch mehr Unterstützung benötigt**. Um dieses Problem anzugehen, ist die Konzeptionierung eines

generischen **MVP** eines DT eine vielversprechende Möglichkeit. Der Hauptfokus der Analyse lag darin, die generischen technischen Bausteine von DTM zu identifizieren. In der Online-Befragung gaben acht von 19 Pilotprojekten an, dass eine solche **generische technische Infrastruktur** für sie vorstellbar wäre.

Es wurden mehrere grundlegende technischen Bausteine der DTM in der Literaturanalyse betrachtet. Dazu zählen Datenbank- oder Speichersysteme, Identitäts- und Zugriffsmanagementsysteme, Schnittstelle zum Datenaustausch und Datensicherheitstechnologien (z.B. Verschlüsselung, Anonymisierung und Pseudonymisierung). In der Online-Befragung haben die Pilotprojekte angegeben, welche dieser technischen Bausteine für das DTM notwendig sind. Im Ergebnis wurden alle Bausteine von mindestens der Hälfte der Pilotprojekte gewählt, aber keinBaustein wurde von allen Projekten befürwortet.

Es lässt sich feststellen, dass alle oben genannten technischen Bausteine in verschiedenen Anwendungsszenarien eine unverzichtbare Rolle spielen, aber es gibt keine Bausteine, die eine universelle Notwendigkeit darstellen. Nichtsdestotrotz haben mehrere Pilotprojekte in Interviews ihr Verständnis von einem Minimum Viable Product (MVP) erwähnt. Das Datenbanksystem und die Datenübertragungsschnittstelle wurden am häufigsten genannt. Zudem teilten zwei befragte Experten auch ihr Wissen über MVP. Einer von ihnen betonte die Bedeutung des Datenbanksystems und der Zugangskontrollkomponente, während der andere Experte der Meinung war, dass das DTM zumindest einen Metadatenkatalog sowie eine BeNutzendeverwaltungskomponente bieten sollte.

Auf der Grundlage dieser Informationen und der Ergebnisse der Literaturanalyse wurde auf dem Workshop ein mögliches MVP vorgestellt, welches **ein Datenverschlüsselungsverfahren, einen Metadatenkatalog oder eine -datenbank, ein Nutzendeverwaltungssystem und ein Zugriffsverwaltungssystem** beinhaltet. Im Zusammenhang damit wurden die Ermöglichung sicherer Datenübertragung, Sicherstellung der Dateninteroperabilität und Gewährleistung der Datensicherheit und -souveränität als die Hauptanforderungen der DTM dargestellt. Dieser Ansatz wurde im Fachgruppenworkshop kontrovers diskutiert. Einige Vertreter von Pilotprojekten argumentierten, dass die Interoperabilität eher als eine optionale Funktion betrachtet werden sollte, anstatt als erforderlich. Zudem wurde die Meinung vertreten, dass ein Nutzendeverwaltungssystem nicht notwendig sei, denn z.B. in einem vollständig anonymisierten System ist die Speicherung von Nutzendedaten sogar nicht erlaubt. In Bezug auf medizinische Anwendungsfälle wurde ein Konsens darüber erzielt, dass das Anonymisierungsverfahren unerlässlich ist. Die Teilnehmenden waren sich am Ende einig, dass die **Konzeptionierung eines MVP nicht umsetzbar ist**. Der Grund liegt darin, dass die Anwendungsfälle und die daraus resultierenden Anforderungen zu unterschiedlich seien. Selbst wenn ein MVP auf der Grundlage eines oder einiger DTM zusammengefasst wird, gibt es keine Garantie, dass das MVP auf alle Anwendungsfälle anwendbar ist. Stattdessen wurde vorgeschlagen, einen **Katalog generischer Herausforderungen** zu entwickeln und die **wiederverwendbaren technischen Bausteine** einzelnen Herausforderungen zuzuordnen. Dieser Vorschlag stieß bei den Pilotprojekten auf positive Resonanz. Ein Pilotprojekt wies darauf hin, dass die Erstellung dieses Katalogs zu ihren Projektzielen passt. Daher wird dieser Katalog in den kommenden AP der Begleitstudie erarbeitet und in späteren Berichten oder auf der Austauschplattform zur Verfügung gestellt.

Allerdings sind die Identifizierung und Zuordnung der wiederverwendbaren technischen Bausteine anspruchsvoll. Die **Zusammenwirkung mit den rechtlichen Rahmenbedingungen** stellt eine der größten Herausforderungen dar. 14 von 19 Pilotprojekte glauben, dass die rechtlichen Anforderungen einen signifikanten Einfluss auf die Entwicklung der technischen Komponenten haben werden. Dazu hat ein Pilotprojekt in der Online-Befragung angemerkt,

dass die unterschiedlichen Anwendungsbereiche jeweils unterschiedliche rechtliche Rahmenbedingungen aufweisen, was zu einer Differenzierung der technischen Bausteine führt. Eine befragte Expertin wies zudem darauf hin, dass die Hauptverantwortung eines DT eher in der Einhaltung **rechtlicher Anforderungen** liegt, wobei die verschiedenen technischen Maßnahmen lediglich unterstützend wirken können. Aufgrund dieser rechtlichen Limitierung stellen technische Bausteine, die in einem oder wenigen Anwendungsbereichen zusammengefasst sind und eine gemeinsame Anforderung, wie die Sicherstellung der Datenqualität, adressieren, die Portabilität ebenfalls nicht sicher. Daher ist es auch bei der Zusammenstellung eines Katalogs wiederverwendbarer technischer Bausteine entscheidend, zwischen spezifischen Anwendungsszenarien zu unterscheiden. Der obige Katalog der technischen Komponenten muss **nach spezifischen Anwendungsszenarien weiter untergliedert** werden, insbesondere dann, wenn in einem Bereich bestimmte rechtliche Anforderungen bestehen.

Zusammenfassend wurde zuerst die Notwendigkeit identifiziert, dass die Datentreuhandstellen in frühen Entwicklungsphasen verstärkt unterstützt werden müssen. Als Lösungsansatz wurde die Konzeption eines MVP für DT ins Auge gefasst, wobei der Fokus auf der Identifizierung **generischer technischer Bausteine** für DTM lag. Generell zeigten die Pilotprojekte Interesse an solcher generischen technischen Infrastruktur. Allerdings wurde während der Interviews mit Pilotprojekten und Experten festgestellt, dass die DT aus verschiedenen Branchen sehr unterschiedliche Anforderungen haben. Von daher haben die Teilnehmenden an dem Fachgruppeworkshop sich geeinigt, dass ein generisches **MVP des DTM nicht umsetzbar** ist. Stattdessen wurde vorgeschlagen, die **wiederverwendbaren technischen Bausteine** je nach Anforderungen zusammenzufassen. Bei der Klassifizierung der Bausteine muss man besonders auf die verschiedenen rechtlichen Anforderungen unterschiedlicher Branchen aufpassen.

5.3 Interoperabilität und Datenübertragung

Das Wert-Risiko-Dilemma wurde in der Literaturanalyse von AP1.1 vorgestellt. Dementsprechend soll der Datentreuhänder sowohl die Risiken des Datenaustauschs minimieren als auch die Wertschöpfung des Datenaustauschs ermöglichen. In diesem Zusammenhang spielt die Interoperabilität eine entscheidende Rolle, um sicherzustellen, dass die ausgetauschten Daten nicht nur zugänglich, sondern auch nutzbar sind. Die Anforderungen an die Interoperabilität sind daher von grundlegender Bedeutung für die erfolgreiche Bewältigung des Wert-Risiko-Dilemmas, dem sich die DT gegenübersehen. In diesem Kapitel werden die Anforderungen und die Herausforderungen bezüglich Interoperabilität und Datenübertragung des DTM anhand der Begleitstudie in AP1.2 vorgestellt.

5.3.1 Anforderungen an die Interoperabilität und Datenübertragung

In den Experteninterviews wurden verschiedene Anforderungen für die Sicherstellung der Interoperabilität im Kontext von DT genannt. Zunächst ist die **Einführung gemeinsamer Schnittstellen und standardisierter Datenformate** von entscheidender Bedeutung. Dies ermöglicht die nahtlose Kommunikation zwischen verschiedenen Datenanbietern und ihren heterogenen Datensätzen. Bei der Erschaffung des Data Space sollten einheitliche Schnittstellen und Datenformate definiert werden, die für alle Teilnehmenden verbindlich sind. Wer Teil des Data Spaces werden möchte, muss sich an diese Vorgaben anpassen und der Nutzung der festgelegten Schnittstellen und Datenformate zustimmen. Eine klare **Definition der Application Programming Interfaces (API) und Datenmodelle** ist ein weiterer essenzieller Schritt, um die Interoperabilität zu gewährleisten. Gleichzeitig sollte das **Authorization Framework auf diese API zugeschnitten** sein, um die Datensicherheit und den Zugriff angemessen zu regeln.

Die individuellen Anforderungen eines Use Cases sollten die Datenströme bestimmen. Inhaltliche Standardisierung und die Verwendung **einheitlicher Ontologien** sind ebenfalls notwendig, um eine kohärente Datenstruktur zu gewährleisten. Diese Ontologien sollten erweiterbar sein und regelmäßig evaluiert werden, um die sich wandelnden Anforderungen berücksichtigen zu können.

In den Projektinterviews wurde des Weiteren darauf hingewiesen, dass die **Verantwortung für die Sicherstellung der Kompatibilität vom konkreten Anwendungsfall abhängt**. Dies erfordert die Definition und Implementierung von Standards, die in der Regel **branchenabhängig** sind. Dabei liegt die Heterogenität der Daten oft in der **Ontologie**, sowohl für Metadaten als auch Nutzdaten. Diese Heterogenität kann mithilfe von **semantischen Techniken auf Schemaebene** erfolgreich adressiert werden. Dies unterstreicht die Bedeutung einer klaren Definition von Standards und Ontologien, um die Interoperabilität zu gewährleisten.

In der Online-Befragung erhielt die Antwort „Standardisierte Schnittstelle zum Datenaustausch" zu der Frage „Welche der technischen Bausteine sind für den Aufbau eines MVP (Minimum Viable Product) Ihres (anvisierten) Datentreuhänders notwendig?" die zweithöchste Zustimmung nach der Antwort „Identitäts- und Zugriffsmanagement-Systeme". Dieses Ergebnis bestätigt die Bedeutung standardisierter Schnittstellen.

5.3.2 Umsetzungshemmnisse bei der Interoperabilität und Datenübertragung

In den Experteninterviews wurden mehrere Umsetzungshemmnisse genannt, die es trotz klarer Anforderungen gibt und welche die Interoperabilität im Kontext von DT beeinträchtigen können. Der **hohe Aufwand** ist eines der zentralen Hindernisse. Es erfordert oft jahrelange Arbeit, um Daten in einen standardisierten Zustand zu überführen. Ein weiteres Problem besteht darin, dass die **Standardisierung der Daten** in einigen Fällen den Schutz sensibler Informationen gefährden kann, insbesondere bei hochsensiblen Daten. Jeder Use Case kann **individuelle Anforderungen** hinsichtlich der zu teilenden Datenformate haben, was die Interoperabilität erschwert. DT sind zudem selten Universalanbieter, sondern entstehen häufig aus bestehenden Systemen oder Beziehungen. **Interessenskonflikte** zwischen den DT und den Datennutzenden können ein weiteres Hindernis für Datenstandards darstellen. Wenn Datennutzende ihre Daten selbst auswerten möchten und eine lokale Datennutzung bevorzugen (im Gesundheitswesen teilweise auch gesetzlich vorgeschrieben), kann dies die Motivation zur Standardisierung der Daten reduzieren und dementsprechend die Komplexität der Interoperabilität erhöhen.

In den Projektinterviews wurden darüber hinaus die **„Sicherstellung der Datenkompatibilität"** sowie der **„Umgang mit Datenheterogenität"** als erhebliche Umsetzungshemmnisse genannt. In vielen Fällen sind sowohl die Anforderungen an die Kompatibilität als auch die Frage des Verantwortlichen für die Sicherstellung der Kompatibilität (z.B. Softwareanbieter, Datennutzende oder DT), branchenspezifisch zu beantworten und vom jeweiligen Anwendungsfall abhängig. Bei heterogenen Daten (sowohl Meta- als auch Nutzdaten) ist der DT in der Verantwortung mit semantischen Techniken die Interoperabilität zu gewährleisten.

Des Weiteren wurde in der Online-Befragung den Projekten die Frage „Welche Hindernisse oder Herausforderungen sind im Laufe des Pilotprojekts aufgetreten bezüglich Interoperabilität und Datenübertragung?" gestellt. Ein zentrales Problem, das von vielen Befragten hervorgehoben wurde, ist der **hohe Aufwand bei der Implementierung einer einheitlichen Kommunikationsschnittstelle**. Dieses Hemmnis haben acht von 19 Befragten ausgewählt und ist damit am häufigsten ausgewählt worden. Außerdem unterstreichen die Umfrageergebnisse, dass die **Harmonisierung unterschiedlicher Datenquellen** und die **Vereinheitlichung heterogener Datenformate und Datenmodelle** als Herausforderung wahrgenommen werden. Beide Aspekte haben sieben von 19 Befragten ausgewählt. Ebenfalls haben sieben von 19

Befragten die Antwort **„Fehlende Dateninfrastrukturen bei den Datengebenden"** ausgewählt. Ohne ausreichende Dateninfrastrukturen auf Seiten der Datenanbieter gestaltet sich der Datenaustausch und die Interoperabilität deutlich schwieriger. Fünf von 19 Befragte wählten auch den **hohen Aufwand bei der Einbindung der existierenden Dateninfrastrukturen** als eine bedeutende Herausforderung. Schließlich wurde die **Bearbeitung großer Datenmengen** von drei und die **fehlenden Informationen bei der Datenfindung** von zwei der 19 Befragten ausgewählt.

5.3.3 Bestehende Datenschnittstellen

In den Projekt- und Experteninterviews wurden verschiedene Schnittstellen genannt, die genutzt werden können, um den Datenaustausch zwischen Datengebenden und Datennutzenden zu ermöglichen. Die Wahl der geeigneten Schnittstellen hängt von den spezifischen Anforderungen des Datentreuhandmodells, den beteiligten Parteien und den Datenschutzbestimmungen ab. **API** ermöglichen die Kommunikation zwischen unterschiedlichen Anwendungen und Systemen. Innerhalb eines DTM können spezielle API entwickelt werden, um den sicheren Datenaustausch zwischen den Datengebenden, dem DT und den Datennutzenden zu ermöglichen. **Standardisierte Kommunikationsprotokolle** können verwendet werden, um den Datenaustausch zwischen verschiedenen Parteien zu erleichtern. In komplexen DTM kann es notwendig sein, **direkte Datenbankverbindungen** herzustellen, um Daten in Echtzeit abzurufen oder zu aktualisieren. Für Datennutzende, die auf die Daten zugreifen und diese bearbeiten möchten, können **Schnittstellen in Form von BeNutzendeoberflächen** entwickelt werden, um die Interaktion zu erleichtern. Außerdem können **Benachrichtigungsschnittstellen** aufgebaut werden, um relevante Informationen, Updates oder Benachrichtigungen z.B. via E-Mail oder App-Benachrichtigung an die beteiligten Parteien zu senden.

Konkrete Beispiele, welche in den Projekten und laut den Experten bereits angewendet werden, sind: Das **Solid-Protokoll** ermöglicht die Verwaltung und gemeinsame Nutzung von Identitätsdaten und persönlichen Informationen innerhalb eines dezentralen Netzwerks. Es handelt sich um eine Webschnittstelle, die es Nutzenden ermöglicht, ihre Identität und Daten sicher zu verwalten und zu teilen. Das **https-Protokoll** wird als Webkommunikationsschnittstelle verwendet, die die sichere Übertragung von Daten zwischen den Datengebenden und Datennutzenden über das https-Protokoll ermöglicht. Die auf Fiware basierte **NGSI-LD-API** ist eine spezifische Schnittstelle, die in Fiware-basierten Lösungen verwendet wird. Sie ermöglicht die Kommunikation und den Datenaustausch innerhalb des Fiware-Ökosystems. Schließlich wurde auch der **GAIA-X international data space connector** genannt, dessen Ansatz auf standardisierte und wiederverwendbare Schnittstellen abzielt. Diese Schnittstellen sind jedoch noch in der Entwicklungsphase.

Dass ein DT die Unterstützung bzw. fertige Komponenten als Schnittstelle für die Datenübertragung anbieten muss, wird durch das Ergebnis aus der Online-Umfrage zu dieser These bestätigt. Sechs von 19 Befragten, die auf diese Frage geantwortet haben, wählten „stimme eher zu" aus und neun von 19 Befragten wählten die Antwort „stimme voll und ganz zu" aus. Die Befragten merkten darüber hinaus an, dass das Fehlen technischer Unterstützung potenzielle Nutzende ausschließen kann, für die die technischen Hürden zu hoch sind. Daher ist es wichtig, fertige Komponenten anzubieten, um die Nutzung des DT zu unterstützen. In der Praxis wäre ein DT ohne solche Komponenten kaum nutzbar, da diese den Datenübertragungsprozess intuitiv und einfach anwendbar machen würden. Außerdem gab es die Anmerkung, dass DT als zentrale Schnittstelle zwischen den Akteuren Komponenten für den Datenaustausch anbieten sollten, um den Abstimmungsaufwand zu minimieren und einen effizienten Daten- und Informationsaustausch zu ermöglichen. Ohne diese Unterstützung wäre

die Zusammenarbeit mit dem DT unwahrscheinlich. Schließlich sollte der DT sicherstellen, dass der Aufwand für die Teilnahme minimal ist. Die Modularität der Komponenten würde es ermöglichen, Funktionalitäten einfach und schnell hinzuzufügen oder zu entfernen, was für eine breite Akzeptanz und Anpassungsfähigkeit des Systems bei den beteiligten Nutzenden entscheidend ist. Die IT-Affinität der Beteiligten spielt dabei eine wichtige Rolle.

5.4 Datensicherheit und -souveränität

DTM müssen sensible Daten in einem vertrauenswürdigen Rahmen verwalten und teilen, was ein hohes Maß an Sicherheit erfordert. Im Folgenden werden die Erkenntnisse aus den Projekt- und Experteninterviews sowie der Online-Umfrage zu den Anforderungen an die Datensicherheit und -souveränität, den Umsetzungshemmnissen und den Lösungsansätzen in den Projekten wiedergegeben.

5.4.1 Anforderungen an Datensicherheit und -souveränität

Die Ergebnisse der Online-Umfrage liefern wichtige Einblicke in die Anforderungen an Datensicherheit und -souveränität in DTM. Diese Umfrageergebnisse geben die Ansichten und Präferenzen der Forschungsprojekte wieder.

In Bezug auf den Aufbau eines MVP für einen DT wurden die Teilnehmenden nach den notwendigen technischen Bausteinen befragt. Die Umfrage zeigt, dass die meisten Befragten **Identitäts- und Zugriffsmanagement-Systeme** (16 von 19), **Bausteine zur Anonymisierung und Pseudonymisierung** (13 von 19) sowie **Datenverschlüsselungstechnologien** (zehn von 19) als unerlässlich für einen DT erachten. Darüber hinaus wurden Anmerkungen hinzugefügt, die die Bedeutung von **Verwendungszweckkontrolle** und einer **neutralen Bewertungsinstanz** (z.B. eine Art Kommission) betonen.

Ein weiterer zentraler Aspekt in DTM betrifft die Datenübertragungskette und ihr Potential in Bezug auf Akzeptanz und Datensicherheit. Die Befragung zeigt, dass die Teilnehmenden unterschiedliche Meinungen dazu haben. Ein Großteil präferiert das Modell **„Datenfluss Peer-to-Peer"** (fünf von 19), bei dem die Daten direkt zwischen den Akteuren ausgetauscht werden. Eine Alternative ist das **„Algorithm to the data"**-Modell (drei von 19), bei dem der DT nur das Ergebnis einer Auswertung an die Datennutzenden weitergibt. Angemerkt wird auch eine Kombination aus „Datenfluss Peer-to-Peer" und „Algorithm to the data". Das Modell **„Questions to the data"**, bei dem Datennutzende eine Anfrage an den DT stellen, welcher diese Anfrage an die Datengebenden weiterleitet und dann für die Weiterleitung der Daten an die Datennutzenden zuständig ist, wurde nur von einem der 19 Befragten angegeben.

Die Frage nach Maßnahmen zur Vermeidung unbefugter Datenübermittlung an Dritte ergab, dass die **Verschlüsselung der Daten während der Übertragung** von 14 der 19 Befragten als unerlässlich erachtet wird. Beinahe ebenso wichtig **sind Nutzendezertifikate und Zugriffskontrollen** für 13 von 19 Befragten. Die Verschlüsselung der gespeicherten Daten sowie eine direkte Verbindung zwischen Datennutzenden und -gebenden während der Datenübertragung als zusätzliche Schutzmaßnahme wurden dagegen nur von fünf bzw. drei von 19 Befragten angegeben.

Zu der Frage, wie sensible Daten in Datentreuhandmodellen am besten geschützt werden, zeigen die Ergebnisse der Umfrage, dass die Teilnehmenden verschiedene Ansätze als wirksam erachten. Zwölf von 19 Befragten sehen die **Pseudonymisierung** der Daten als wichtige Maßnahme, um Rückschlüsse auf den Datengebenden zu verhindern. Zusätzlich ist für neun von 19 Befragten die **asymmetrische Kryptographie** von Bedeutung, um sicherzustellen, dass keine unbefugten Dritten, einschließlich des DT, Zugriff auf die Daten haben. Darüber hinaus befürworten 13 von 19 Befragten **mehrstufige Zugriffs- und Nutzungskontrollen**, die sicherstellen,

dass nur autorisierte Datennutzende auf die Daten zugreifen dürfen und diese nur für erlaubte Zwecke nutzen können. In den Anmerkungen haben die Teilnehmenden darüber hinaus auf Datensparsamkeit und selektive Offenlegung hingewiesen.

In Bezug auf Verfahren für die Gewährleistung der Sicherheit und Zuverlässigkeit bei der Prüfung von Datenzugriffsrechten war die Meinung geteilt. Sieben von 19 Befragten präferierten, dass **die Datengebenden die Datenzugriffsrechte prüfen** müssen. Die Befragten sehen die **Verantwortung** allerdings mehr **beim DT**, entweder mit einer **zentralen Komponente** (acht von 19) oder durch **Bereitstellung einer dezentralen Komponente** für Datengebende (neun von 19) oder Datennutzende (vier von 19). Einige Teilnehmende betonten, dass die verschiedenen Verfahren je nach den spezifischen Anforderungen und Use Cases des Projekts angewendet werden können. In den Anmerkungen wurde hervorgehoben, dass die **Kombination verschiedener Verfahren** zur Prüfung der Datenzugriffsrechte wahrscheinlich die beste Möglichkeit darstellt, um Sicherheit und Zuverlässigkeit zu gewährleisten. Außerdem wurde angemerkt, dass grundsätzlich die Datengebenden entscheiden müssen, wer Zugriff auf ihre Daten erhält. Demzufolge hat der **DT die Aufgabe, nach den Vorgaben der Datengebenden die Zugriffsrechte zu vergeben** und diese bei Datenanfragen zu überprüfen.

5.4.2 Hemmnisse bei der technischen Umsetzung zur Sicherstellung der Datensicherheit und -souveränität

In den Projekt- und Experteninterviews wurden verschiedene Umsetzungshemmnisse bezüglich Datensicherheit und -souveränität herausgestellt, mit denen sich die Projekte aktuell beschäftigen. Ein Aspekt ist die **Balance zwischen Verschlüsselung und Durchsuchbarkeit**. Der DT trägt die Verantwortung dafür, Metadaten so zu handhaben, dass die Datennutzenden die Daten für gewünschte Analysen leichter finden können, ohne dass dabei zu viele Informationen über die Datengebenden preisgegeben werden. Eine weitere Herausforderung besteht in der Frage, inwieweit **Metadaten untereinander verknüpft** und einem bestimmten Datengebenden zugeordnet werden können. Ein bedeutendes Problem besteht, sobald ein Datennutzende **physischen Zugriff auf die Daten** hat. Technisch kann dann die weitere Nutzung der Daten nicht mehr eingeschränkt werden, was die Kontrolle über die Datensicherheit erschwert. Die Sicherstellung des **Widerrufsrechts der Datengebenden** ist ebenfalls von großer Bedeutung. Die Löschung von Daten darf nicht dazu führen, dass Rückschlüsse auf den Ursprung der Daten und die Identität der Teilnehmenden gezogen werden können. Schließlich stehen viele Datentreuhandstellen vor der **Herausforderung, angemessene technische Komponenten zur Gewährleistung der Datensicherheit zu identifizieren und zu implementieren**. Dies kann dazu führen, dass Sicherheitslücken übersehen und Risiken unterschätzt werden, wodurch erhebliche potenzielle Schadenssummen entstehen können. Es besteht ein dringender Bedarf an einheitlichen technischen Standards für Sicherheitsmaßnahmen, die eine effektive Überprüfung der Datensicherheit ermöglichen und beispielsweise die Ausstellung von Sicherheitszertifikaten erleichtern. Die Etablierung dieser Standards ist von entscheidender Bedeutung, um die Integrität und Vertraulichkeit der Daten in DTM zu gewährleisten.

Die Ergebnisse der Online-Umfrage unterstreichen und ergänzen die in den Interviews genannten Punkte. Die Befragten gaben verschiedene **Hindernisse und Herausforderungen bezüglich Datensicherheit und -souveränität** an. Ein wichtiger Aspekt war das Fehlen entscheidender technischer Bausteine, die zur Authentifizierung (drei von 19), Autorisierung (5fünfvon 19) und Nutzungskontrolle (acht von 19) beitragen. Diese Bausteine sind entscheidend, um sicherzustellen, dass nur autorisierte BeNutzende auf die Daten zugreifen können und deren Verwendung kontrolliert wird. Des Weiteren zeigt sich, dass technische Maßnahmen zur Datenanonymisierung (neun von 19) und Pseudonymisierung (sieben von 19) von großer Bedeutung sind, um die Privatsphäre der Datengebenden zu schützen und

Rückschlüsse auf individuelle Identitäten zu verhindern. Der Schutz sensibler Daten vor externem Zugriff (sechs von 19) und unbefugter Nutzung (sieben von 19) wurde ebenfalls als Herausforderung identifiziert, was die Notwendigkeit von robusten Sicherheitsmechanismen betont, um die Integrität und Vertraulichkeit der Daten zu gewährleisten. Schließlich wird die Sicherstellung des Widerrufsrechts von geteilten Datensätzen (zehn von 19) als kritisch angesehen, um Datensouveränität und die Kontrolle über die eigenen Daten zu gewährleisten. Dies unterstreicht die Bedeutung von technischen Maßnahmen, die das Vertrauen in die Datensicherheit und -souveränität im DTM gewährleisten.

5.4.3 In der Praxis genutzte technische Maßnahmen

In den Projektinterviews wurden von den Projekten verschiedene Maßnahmen genannt, die Datensicherheit und -souveränität sicherstellen und innerhalb der Projektarbeiten umgesetzt werden. Bevor Daten übertragen werden, sind strenge **Autorisierungs- und Authentifizierungsprozesse** erforderlich. Diese stellen sicher, dass nur berechtigte Nutzende auf die Daten zugreifen können, und dass die Identität der Datengebenden und -nutzenden verifiziert wird. Ein **effizientes API-Design für die Datengebenden** muss sicherstellen, dass nur spezifische Daten freigegeben werden, die für den jeweiligen Zweck benötigt werden. Ein schlechtes API-Design kann zu unnötiger Offenlegung von Daten führen. Dies ermöglicht einen präzisen, isolierten Zugriff auf die benötigten Daten. Die **Datenverschlüsselung** erfolgt beispielsweise mithilfe des AES. Der DT erhält keinen Schlüssel, um die Daten zu entschlüsseln. Die sichere Schlüsselfreigabe wird durch **asymmetrische Kryptographie** gewährleistet. Der Datengebende verschlüsselt den Datenfreigabeschlüssel mit dem öffentlichen Schlüssel des Datennutzenden, um sicherzustellen, dass der DT keine Einblicke in die Schlüsselübergabe hat. Zusätzliche Schutzmaßnahmen können auf den Endgeräten der Nutzenden durch die **Datenpseudonymisierung** durchgeführt werden. Dies hilft, die Identifizierung über die Daten zu verhindern. Schließlich findet **Datenverschlüsselung** während der Übertragung innerhalb des Netzwerks statt, um die Vertraulichkeit zu gewährleisten. Ein **Policy Monitor oder Protokolldateien** bieten die Möglichkeit, den Datenfluss zu überwachen und bei Missbrauch oder Fehlverhalten einzugreifen.

Die befragten Experten gaben in den Interviews darüber hinaus an, welche weiteren Maßnahmen ihnen aus der Praxis bekannt sind. Zu diesen Maßnahmen zählt die **föderierte Datenhaltung**, bei der Daten in dezentralen Systemen gespeichert werden, um das Risiko eines Single Points of Failure zu minimieren. Ein **hierarchisches oder rollenbasiertes Zugriffssystem**, wird implementiert, um Zugriffsrechte und -umfang zu definieren. Dies ermöglicht eine präzise Steuerung des Datenzugriffs, einschließlich Zugriff nur auf bestimmte Attribute oder die Erlaubnis zum Senden von Daten. Des Weiteren müssen **IT-Sicherheitsrichtlinien** implementiert werden, die als Grundlage für den sicheren Datenumgang in einem DTM dienen. Schließlich müssen alle Systemebenen immer nach dem aktuellen Stand der Technik abgesichert werden.

5.5 Zusammenfassung und Ausblick

In AP1.2 wurde zuerst eine Analyse der Vorhabensbeschreibungen sowie der Zwischenberichte durchgeführt. Die Pilotprojekte weisen unterschiedliche Schwerpunkte auf, wobei die Mehrheit auf bestehende technische Ansätze zurückgreift und nur knapp 30 % Grundlageforschung oder Neuentwicklungen betreiben. Die Pilotprojekte haben weitgehend die Anforderungsanalyse und Literaturrecherche abgeschlossen und teilweise mit der prototypischen Umsetzung begonnen.

Hinsichtlich der **technischen Architekturen und Bausteine** wurde eine Vielzahl von Ansätzen wahrgenommen. Darunter sind Data Space und Datenintermediär die am häufigsten erwähnten Ansätze. Allerdings kann die Vielzahl an verfügbaren Architekturen und

technischen Bausteinen die Auswahl für neue DT erschweren und sogar zu ungeeigneten technischen Entscheidungen führen. Die Zusammenfassung von generischen oder wiederverwendbaren technischen Bausteinen wurde als Lösung in Betracht gezogen. Es zeigte sich jedoch im Verlauf der Begleitstudie, dass die Identifizierung der generischen technischen Bausteine nur schwer umsetzbar ist. Die Klassifizierung der wiederverwendbaren technischen Bausteine steht ebenfalls vor großen Herausforderungen, darunter die Berücksichtigung rechtlicher Beschränkungen und branchenspezifischer Anforderungen.

Als weitere Erkenntnis von AP1.2 wurde identifiziert, dass die **Interoperabilität** von DTM einheitliche Schnittstellen, Datenformate und Ontologien erfordert. Die Implementierung stößt jedoch auf Hindernisse durch den hohen Aufwand, Datenschutzbedenken und Interessenskonflikte. Eine klare Definition von Standards und Ontologien ist entscheidend, um Interoperabilität zu gewährleisten. Pilotprojekte und Experteninterviews ergaben, dass DT auf verschiedene Schnittstellen zugreifen können, darunter API, standardisierte Kommunikationsprotokolle und direkte Datenbankverbindungen. Wiederverwendbare Schnittstellen für die Datenübertragung sind von großer Bedeutung, um die Umsetzbarkeit sicherzustellen. Die Modularität der technischen Bausteine unterstützt die Akzeptanz und Anpassungsfähigkeit des Systems.

In der Online-Befragung zur **Datensicherheit und -souveränität** in DTM betonen die Teilnehmenden die Bedeutung von Identitäts- und Zugriffsmanagement, Anonymisierung, Pseudonymisierung und Datenverschlüsselung. Verschiedene Ansichten zur Datenübertragungskette und Prüfung von Datenzugriffsrechten wurden aufgezeigt. Die Bedeutung von technischen Maßnahmen wie Pseudonymisierung, asymmetrischer Kryptographie und mehrstufigen Zugriffs- und Nutzungskontrollen wurde hervorgehoben. Als Umsetzungshemmnisse im Bereich Datensicherheit und -souveränität wurden insbesondere die Abwägung zwischen Verschlüsselung und Durchsuchbarkeit und die Gewährleistung des Widerrufsrechts von Datengebenden angegeben. Die Pilotprojekte haben bereits einige Autorisierungs- und Authentifizierungsprozesse entwickelt. Die Maßnahmen wie AES und Asymmetrische Kryptographie wurden verwendet, um die Datensicherheit zu gewährleisten.

Forschungsbedarf besteht damit insbesondere noch im Bereich des Zusammenwirkens zwischen den technischen Bausteinen und den rechtlichen Rahmenbedingungen. Die Entwicklung gemeinsamer Schnittstellen und standardisierter Datenformate ist für die Nutzbarkeit von DTM entscheidend und damit weiter voranzutreiben. Wie weit der Verantwortungsbereich des DT hinsichtlich der Datenqualität oder einer weitergehenden Datenaufbereitung geht, kann sehr unterschiedlich ausfallen und muss weitergehend im Kontext von möglichen Interessenskonflikten betrachtet werden. Generell erhöhen leicht nutzbare Systeme und zusätzliche Datenaufbereitungskonzepte sowie die Gewährleistung von Sicherheit und Zuverlässigkeit die Attraktivität der Konzepte.

6 QT2: Rechtliche Rahmenbedingungen und Ausgestaltung der DTM

Auch die inhaltliche Schwerpunktsetzung in QT 2 in AP1.2 baut auf den Ergebnissen von AP1.1 auf. Insbesondere werden die in AP1.1 identifizierten Forschungslücken adressiert, die in der übergreifenden Einführung in Kapitel 4 (Tabelle 3) dargestellt wurden. Die inhaltliche Schwerpunktsetzung in QT 2 konzentriert sich in AP1.2 auf die folgenden drei Bereiche:

Rechtliche Herausforderungen für Datenteilende,

Rechtliche Herausforderungen für DT,

Strategien zur Bewältigung der rechtlichen Herausforderungen (insbesondere des Wert-Risiko-Dilemmas)

6.1 Rechtliche Herausforderungen für Datenteilende

Die Bestandsaufnahme aus AP1.1 ergab, dass in der Diskussion bisher noch nicht hinreichend die **Funktionen von DT** herausgearbeitet wurden. Insbesondere bleibt bisher unklar, welche Probleme ein DT eigentlich konkret lösen soll. Oft wird hierbei vom fehlenden Vertrauen der Datengebenden gesprochen, was dazu führe, dass Daten nicht ausreichend geteilt werden. Es bleibt aber oftmals unklar, worauf sich das fehlende Vertrauen der Datenteilenden bezieht bzw. welches Vertrauen ein DT bei den Datenteilenden herstellen oder absichern soll.

6.1.1 Rechtliche Herausforderungen als Bestandteil des Wert-Risiko-Dilemmas

In der Bestandsaufnahme wurde der Fokus dafür auf das sogenannte **Wert-Risiko-Dilemma** gelegt. Demnach werden Daten deshalb nicht geteilt, weil ein Datenhaltender deutlich mehr Compliance Risiken und Kosten als einen Mehrwert im Teilen seiner Daten sieht. Die Funktion von DT wäre demnach, dieses Wert-Risiko-Dilemma bei den Datenteilenden aufzulösen (bzw. zumindest dabei zu helfen). Dieses Konzept ist in der folgenden Abbildung veranschaulicht.

Abbildung 4 Wert-Risiko-Dilemma beim Teilen von Daten

Quelle: Eigene Darstellung (Law & Innovation)

Ein Datenhaltender sieht also zum Zeitpunkt der Datenteilung oftmals nur einen abstrakten Mehrwert, während er mit dem Teilen seiner Daten mehrere konkrete Compliance-Risiken eingeht, allen voran den Datenschutz und/oder die IT-Sicherheit nicht einzuhalten oder Geschäftsgeheimnisse zu verletzen. Dagegen kann der Datennutzende dem Datenhaltenden oftmals noch kein konkretes Mehrwertversprechen machen, weil der Datennutzende in der Regel selbst noch nicht weiß, auf welche Weise er mit den Daten einen Mehrwert generieren kann, an dem er den Datenhaltenden partizipieren lassen könnte. Dies gilt vor allem dann, wenn sich der Mehrwert nicht unmittelbar in einem finanziellen Mehrwert niederschlägt, wie etwa im Bereich der öffentlichen Vorsorge oder bei karitativen Zielen. Auch bei Innovationsprozessen stellt sich das Dilemma, weil diese durch ihre Ergebnisoffenheit geprägt

sind und sich ein konkreter Mehrwert erst im Laufe der Zeit, nicht aber bereits zum Zeitpunkt des Datenteilens herausstellt. Selbst wenn die Datenteilenden den Mehrwert soweit steigern und die Risiken soweit minimieren könnten, dass der Mehrwert höher als die Risiken wäre, sind die Kosten, die durch die technisch-organisatorischen Maßnahmen für die Wertschöpfung und Risikokontrolle anfallen, oftmals prohibitiv hoch. **Die Funktion eines DT ist also, die Datenteilenden durch die Bereitstellung von Diensten dabei zu unterstützen, den Mehrwert so zu steigern sowie die Compliance-Risiken und Kosten so zu senken, dass das Teilen der Daten aus Sicht der Datenteilenden lohnenswert ist.**

Die Klärung der **Funktion von DT** ist nicht zuletzt für die vorliegende Begleitforschung wichtig, weil hier oft sofort von Herausforderungen der DT gesprochen wird, anstatt zunächst von den Herausforderungen der Datenteilenden zu sprechen. Diese müssen aber nicht notwendigerweise dieselben sein. Deshalb wird hier zunächst von den **Herausforderungen der Datenteilenden** gesprochen. Erst dann wird untersucht, welche Herausforderungen die DT haben, die Datenteilenden bei der Überwindung ihrer Herausforderungen zu unterstützen.

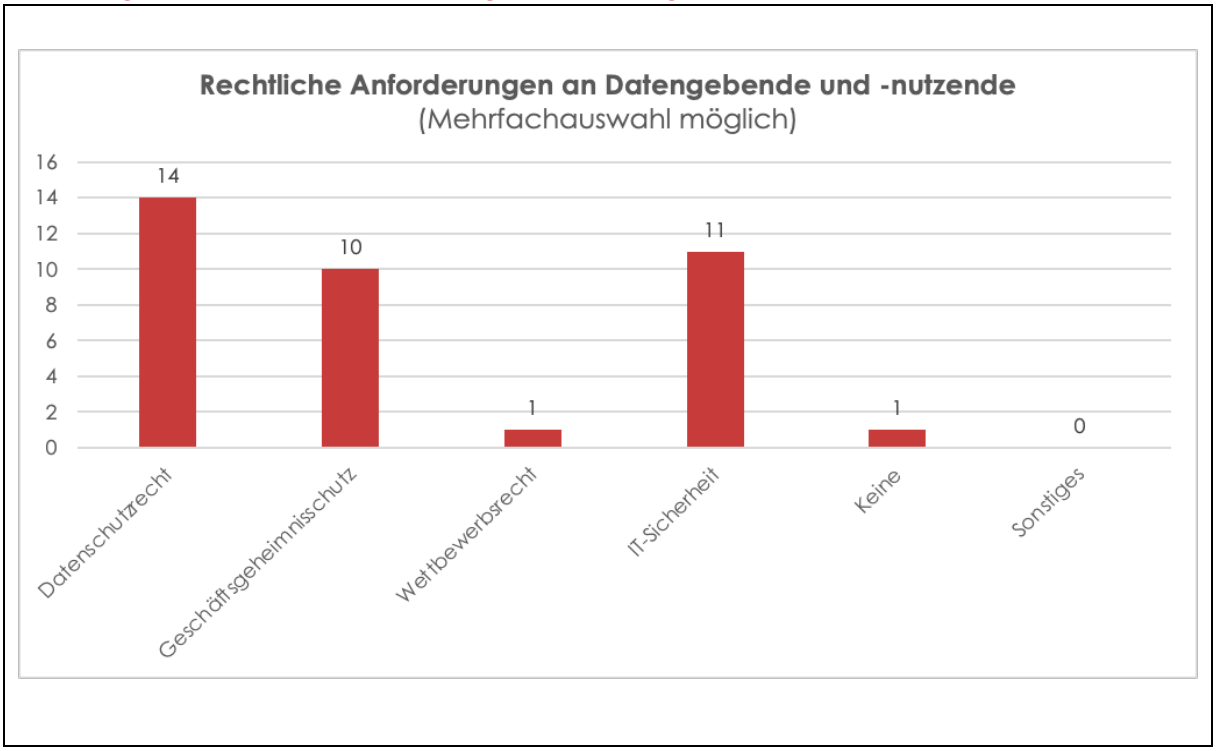
Grundsätzlich kommt das Wert-Risiko-Dilemma nicht nur in Situationen, in denen Daten freiwillig geteilt werden, zum Tragen, sondern auch in solchen, in denen Daten (wie etwa im Data-Act-Entwurf vorgesehen) geteilt werden müssen. Denn selbst wenn der Datennutzende einen gesetzlichen Anspruch auf den Zugang zu Daten hat, muss er hierbei konfligierende Schutzinteressen berücksichtigen, wie etwa den Datenschutz oder Geschäftsgeheimnisse. Der Unterschied zum freiwilligen Teilen der Daten ist also, dass es jetzt nicht mehr auf die Abwägung des Datenhaltenden, sondern des Datennutzenden ankommt. Denn dann stellt sich dieser die Frage, ob es sich lohnt, das gesetzliche Datenzugriffsrecht trotz der damit verbundenen Compliance Risiken und Kosten geltend zu machen. Da bei den vorliegenden Pilotprojekten die Daten aber fast ausschließlich freiwillig geteilt werden (siehe Frage 24 der Online-Survey), wird hier auf Situationen des mandatorischen Datenteilens nicht weiter eingegangen.

Der Vorschlag, dass DT primär die Funktion haben (sollten), die Datenteilenden dabei zu unterstützen, zumindest die **Compliance Risiken und die Kosten des Datenteilens** zu senken, wurde in den Erhebungen für dieses AP weitgehend bestätigt. Ein Experte warf im Rahmen eines Interviews die Frage auf, ob DT zumindest auch die Funktion haben (sollten), auf makroökonomischer Ebene der Übermacht sehr großer datengetriebener Unternehmen entgegenzuwirken. Dagegen wurde in der Diskussion angebracht, ob die Einschränkung der Übermacht solcher Konzerne nicht eher das Ziel anderer EU-Gesetze wäre, wie etwa des Digital Services Act und des Digital Market Act. Demgegenüber verfolgt die Idee von DT das Ziel, Akteure zu unterstützen, die bisher (eben aufgrund der Compliance Risiken und Kosten) nicht in der Lage sind, Daten in gleichem Umfang zu nutzen, wie diese marktmächtigen datengetriebenen Unternehmen. Einigkeit bestand zumindest darin, dass DT die Funktion haben, die Datenteilenden bei der Einhaltung der rechtlichen Vorgaben und mit den nötigen technisch-organisatorischen Maßnahmen zu unterstützen.

6.1.2 Rechtliche Herausforderungen im Einzelnen

Auch in den Interviews mit den Pilotprojekten stellte sich heraus, dass die Datenhaltenden und Datennutzende vor allem rechtliche Hemmnisse beim Datenteilen sehen und dass DT dabei versuchen, diese vor allem mit technisch-organisatorischen Maßnahmen beim Datenteilen zu unterstützen. Diese Beobachtung bestätigt sich auch in dem Online-Survey. Die folgenden Grafiken zeigen, bei welchen rechtlichen Vorgaben und mit welchen Maßnahmen DT die Datenteilenden unterstützen.

Abbildung 5 Rechtliche Anforderungen an Datengebende und -nutzende

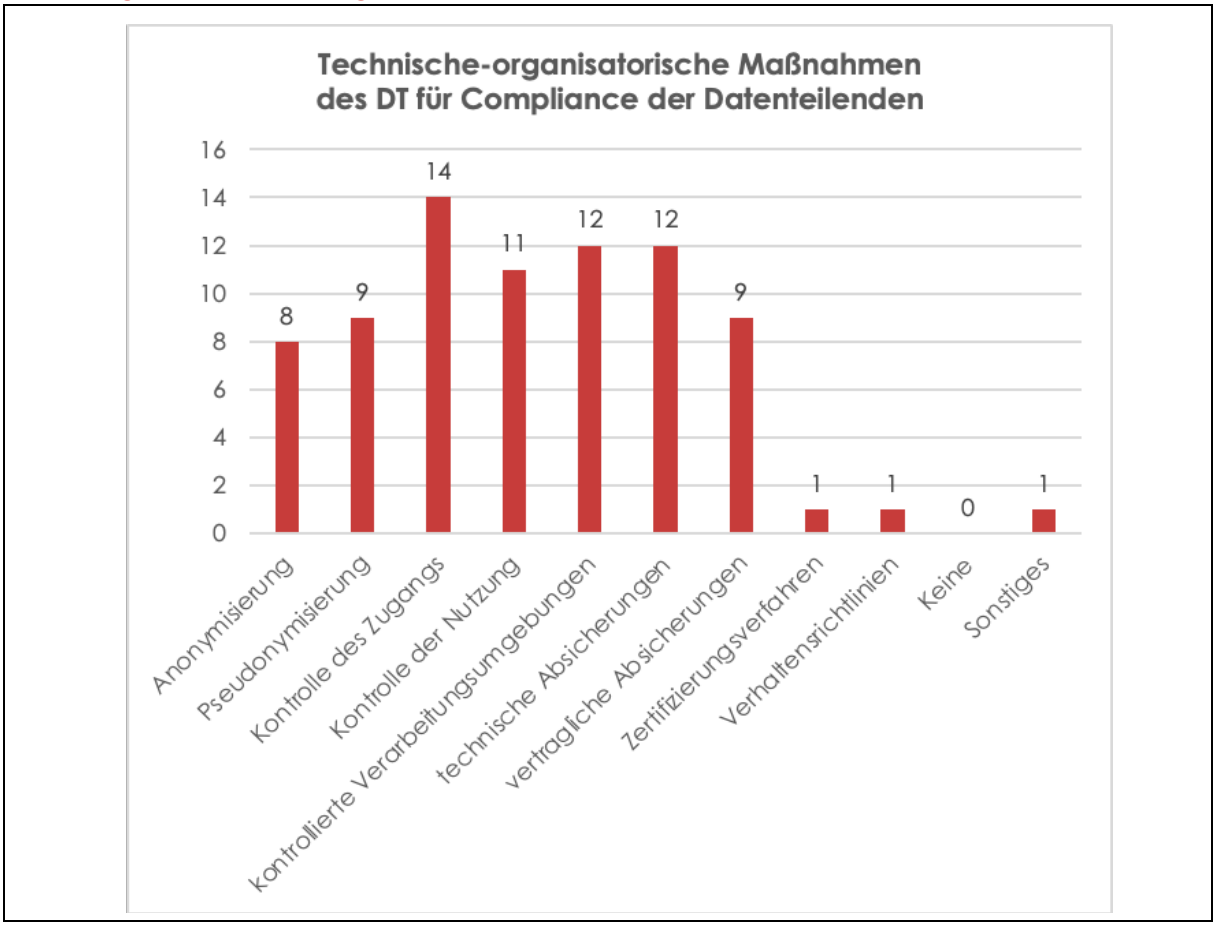


Quelle: Eigene Darstellung (Law & Innovation) auf Grundlage der Online-Befragung der Pilotprojekte (n=18)

Bei den Ergebnissen zu den rechtlichen Anforderungen ist zunächst interessant, dass das **Wettbewerbsrecht** bisher keine große Rolle bei den Überlegungen der DT spielt. Dabei ist es durchaus denkbar, dass der Austausch von Daten zwischen Akteuren in bestimmten Konstellationen als unlautere Absprache im Sinne des Gesetzes gegen Wettbewerbsbeschränkungen angesehen werden kann. Ob solche Konstellationen bei den Pilotprojekten schlicht nicht vorliegen oder ob es sich hier um blinde Flecken in der Wahrnehmung der DT handelt, konnte an dieser Stelle noch nicht geklärt werden.

Daneben ist ein Ergebnis aus dem Fachgruppenworkshop hervorzuheben, das sich bereits in den Interviews mit den Pilotprojekten abzeichnete, aber so klar aus dem Online-Survey nicht hervorgeht: Eine der größten Herausforderungen, mit denen sich die Pilotprojekte konfrontiert sahen, war die Frage, welche der am Teilen der Daten beteiligten Akteure welche technisch-organisatorische Maßnahme genau vornehmen und welche **Haftung** dafür tragen müssen. Auch die Klärung dieser Frage kann ein DT übernehmen. Die folgende Abbildung zeigt, dass die Pilotprojekte diese Frage zumindest implizit beantworten, indem sie selbst zahlreiche der erforderlichen technisch-organisatorischen Maßnahmen erbringen.

Abbildung 6 Technisch-organisatorische Maßnahmen



Quelle: Eigene Darstellung (Law & Innovation) auf Grundlage der Online-Befragung der Pilotprojekte (n=18)

Bei den von den DT bereit gestellten **technisch-organisatorischen Maßnahmen** zur Unterstützung der Datenteilenden sind zwei Beobachtungen interessant. Zum einen scheint die Idee, den Datenteilenden kontrollierte Verarbeitungsumgebungen zur Verfügung zu stellen, bereits weite Verbreitung zu finden. Dahinter steht die Idee, dass ein Datenhaltender einem Datennutzenden die Daten nicht einfach zur freien „Verwendung" übergibt, sondern dass der Datennutzende die Daten nur im Rahmen einer kontrollierten Verarbeitungsumgebung nutzen darf. So kann in Entsprechung der rechtlichen Compliance Risiken, die für den Datenhaltenden durch die Weitergabe der Daten entstehen, mit mehr oder weniger strengen Bedingungen und Kontrollen sichergestellt werden, dass der Datennutzende die Daten nur in risiko-reduzierter Weise verwendet. Die Idee wird zwar seit einigen Jahren in einzelnen Bereichen erfolgreich praktiziert (so etwa im untersuchten externen Use Case der Forschungsdatenzentren des Bundes und der Länder). Auch der Gesetzgeber greift die Idee zunehmend auf (so etwa im Data Governance Act für die Bereitstellung geschützter Daten durch die öffentliche Hand, Art. 5 Abs. 4). Bis dato hatte die Idee allerdings noch keine weite Verbreitung erfahren, weder in der Literatur noch in der Praxis. Bei den Pilotprojekten scheinen kontrollierte Verarbeitungsumgebungen bereits Bestandteil üblicher Praxis zu sein (unten unter Kapitel 4.3 mehr zu den verschiedenen Ausprägungen „kontrollierter Verarbeitungsumgebungen").

Daneben ist die Beobachtung interessant, dass die DTM für die Datenteilenden bisher kaum **Zertifizierungsverfahren** oder **Verhaltensrichtlinien** bereitstellen. Das ist deshalb interessant, weil solche Verfahren besonders geeignet sind, die hohe Rechtsunsicherheit zu beseitigen. In der

DSGVO sind solche Verfahren vorgesehen (Art. 40 ff.), damit Datenverarbeiter in ihrem konkreten Fall die Einhaltung der DSGVO nachweisen können (genauer gesagt kann die Einhaltung von Verhaltensrichtlinien oder eines Zertifizierungsverfahrens beim Nachweis der DSGVO-Konformität als ein Gesichtspunkt herangezogen werden, siehe insbesondere Art. 24 Abs. 3, Art. 25 Abs. 3, Art. 32 Abs. 3). Zur Bewältigung der Compliance Risiken benötigen Datenteilende also nicht nur die technisch-organisatorischen Maßnahmen, die für die Kontrolle der Risiken nötig sind, sondern auch Mechanismen, mit denen sie die Konformität nachweisen können. Solche Mechanismen bieten die Pilotprojekte bisher kaum an. Damit stellt sich die Frage, ob diese Verfahren von den Pilotprojekten übersehen werden. In den Fachgruppenworkshops ergab sich allerdings ein differenzierendes Bild. Denn die Pilotprojekte scheinen zunächst zwischen den Compliance Risiken zu unterscheiden:

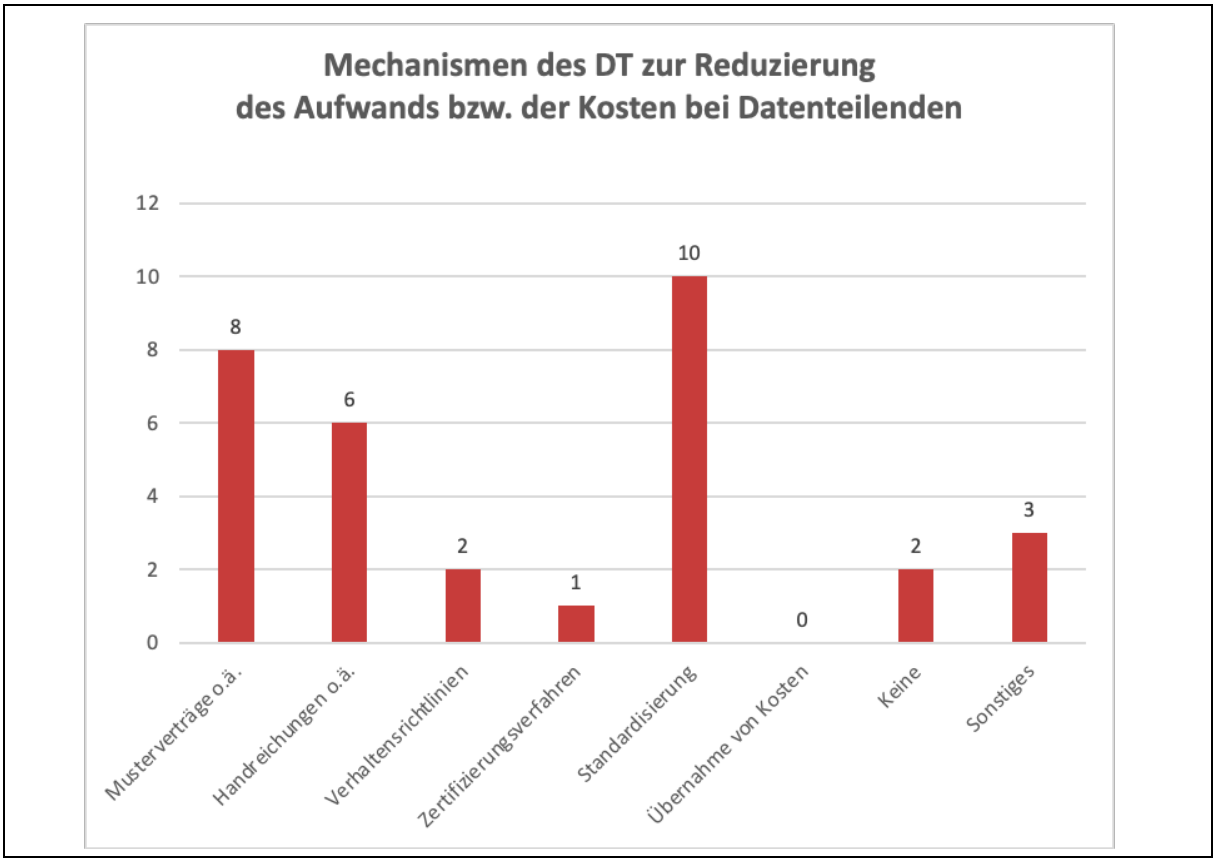
Soweit es sich um das **Compliance Risiko** handelt, dass die Datenteilenden ein Geschäftsgeheimnis verletzen, scheinen sie auf die vom DT angebotene technisch und/oder organisatorische Lösung zu vertrauen, ohne dass ein weiterer, gesetzlich verankerter bzw. unterstützter Nachweis der Rechtskonformität nötig wäre. Geht es dagegen um den Nachweis der Konformität mit dem Datenschutzrecht, besteht dagegen ein erhöhtes Bedürfnis nach Rechtssicherheit. Der Grund hierfür ist offenbar, dass eine Verletzung von Geschäftsgeheimnissen zivilrechtlich zwischen den Parteien geregelt wird, während der Vollzug des Datenschutzrechts eben auch von den Datenschutzbehörden hoheitlich ausgeführt wird. Hierfür reicht es gerade nicht, dass die Betroffenen die Rechtsfolgen im Wege eines DT zivilrechtlich regeln.

Aber auch beim Thema **Datenschutzrecht** scheinen die Datenteilenden nicht immer einen gesetzlich formalisiertes Nachweisverfahren für ihre Rechtskonformität für notwendig zu halten. Aus den Fachgruppenworkshops ergab sich vielmehr, dass wenn sich DT ihrerseits auf Handreichungen der Datenschutzbehörden oder ähnliche Standards berufen können, auch das als ausreichender Vertrauensanker bzw. ausreichendes Signal der Rechtskonformität angesehen werden kann.

6.1.3 Mechanismen zur Reduzierung des rechtlichen Compliance Aufwands

Damit sind wir schließlich bei den Mechanismen, die DT zur Senkung des Aufwands bzw. der Kosten bei den Datenteilenden einsetzen. Hintergrund ist hierbei, dass der Aufwand oftmals unverhältnismäßig hoch ist, der für die Klärung der mit dem Teilen von Daten verbundenen Rechtsfragen verbunden ist. Um diese Kosten bei den Datenteilenden zu senken, das heißt aufwändige und kostenintensive rechtliche Einzelfallprüfungen zu vermeiden, setzen DT laut den Ergebnissen der Online-Studie folgende Instrumente ein:

Abbildung 7 Mechanismen des DT zur Reduzierung des Aufwands bzw. der Kosten bei Datenteilenden



Quelle: Eigene Darstellung (Law & Innovation) auf Grundlage der Online-Befragung der Pilotprojekte (n=18)

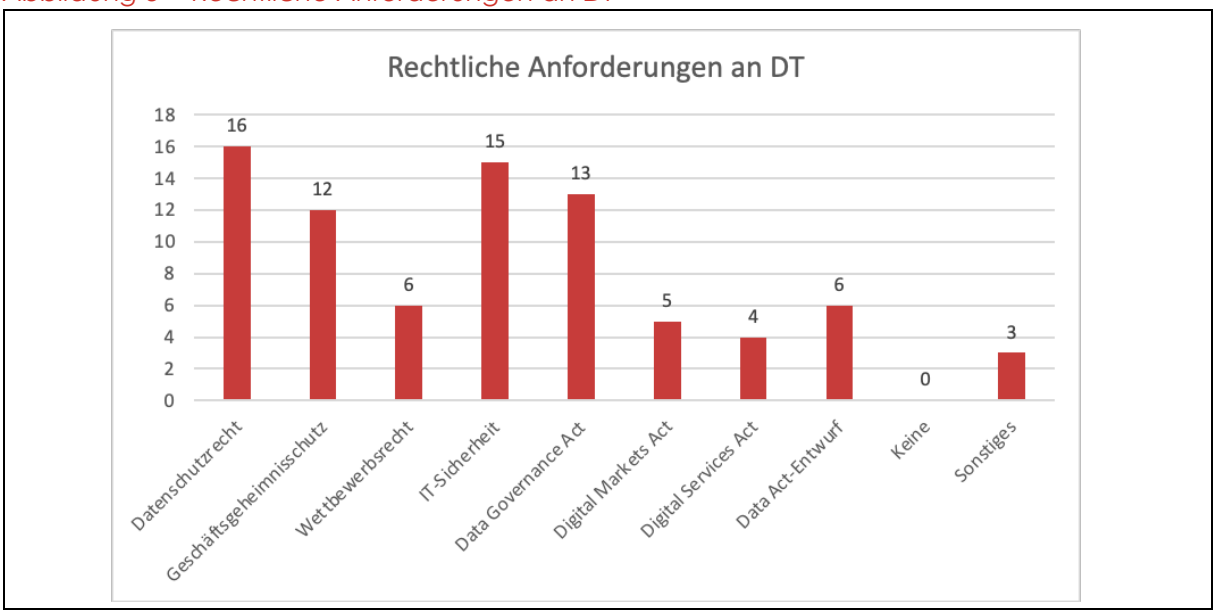
Noch vor Musterverträgen und Handreichungen spielen **Standardisierungen**, wie bestimmte rechtliche Vorgaben technisch-organisatorisch konkret umgesetzt werden müssen, bei den Pilotprojekten die wichtigste Rolle. Hier gibt es eine deutliche Schnittmenge zu den Skalierungspotentialen, wie in QT 4 beschrieben. Es fällt auf, dass auch hier wieder Verhaltensrichtlinien und Zertifizierungsverfahren, wie sie etwa in Art. 40 ff. DSGVO vorgesehen sind, nur eine untergeordnete Rolle spielen. Ein Grund ist, dass sich die Entwicklung von Verhaltensrichtlinien und Zertifizierungsprogrammen im Zusammenspiel mit der Deutschen Akkreditierungsstelle (DAkkS) und den zuständigen Datenschutzbehörden als so aufwändig herausgestellt hat, dass dies von kaum einem DT mit den begrenzten Ressourcen eines KMU geleistet werden kann. Selbst wenn es einem DT oder einem anderen KMU gelänge, eine entsprechende Richtlinie oder ein Zertifizierungsprogramm von den Behörden erfolgreich akkreditieren zu lassen, stellte sich immer noch die Anschlussfrage, ob Datenteilende das ebenfalls als aufwändig bekannte Verfahren zur Unterwerfung unter eine Richtlinie oder eine Zertifizierung als lohnenswert ansehen. In Ansehung der oben beschriebenen Tatsache, dass diesen oftmals die Anwendung von Handreichungen oder anderen Standardisierungen reichen, um die Rechtsunsicherheit zu reduzieren, darf dies bezweifelt werden.

6.2 Rechtliche Herausforderungen für DT

Auch wenn hier zwischen den rechtlichen Herausforderungen für die Datenteilenden und den rechtlichen Herausforderungen für die DT unterschieden wird, muss das nicht bedeuten, dass

diese in jedem Fall verschieden sein müssen. Um das aber feststellen zu können, muss zunächst zwischen den Akteuren unterschieden werden. Tatsächlich liegt es nahe, dass sich die rechtlichen Herausforderungen der Datenteilenden bei den DT widerspiegeln. Denn sobald Datenteilende auf DT für die Bewältigung ihrer rechtlichen Herausforderungen zurückgreifen, stellt sich auch für DT die Frage, wer welche Maßnahmen rechtlich erbringen darf bzw. muss und wer dafür jeweils haftet. Die folgende Grafik zeigt die Ergebnisse aus der Online-Studie, in der die Pilotprojekte gefragt wurden, welche rechtlichen Anforderungen sie selbst einhalten müssen.

Abbildung 8 Rechtliche Anforderungen an DT



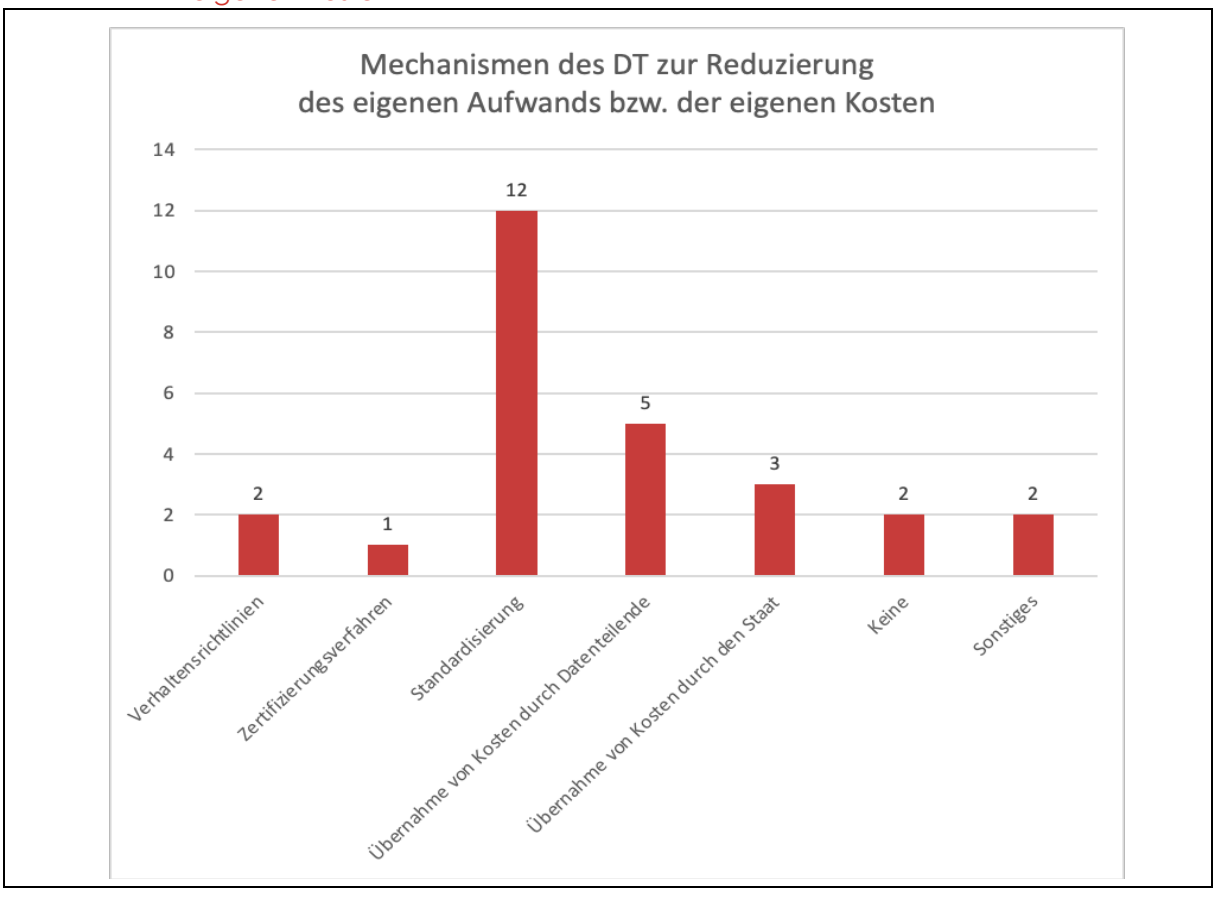
Quelle: Eigene Darstellung (Law & Innovation) auf Grundlage der Online-Befragung der Pilotprojekte (n=18)

Die Ergebnisse sind im Vergleich zu den genannten rechtlichen Herausforderungen für die Datenteilenden in dreierlei Hinsicht aufschlussreich. Zunächst fällt auf, dass die Einhaltung des Datenschutzrechts, der IT-Sicherheit und die Wahrung der Geschäftsgeheimnisse auch für die DT die wichtigsten Herausforderungen darstellen. Darüber hinaus wird nun aber auch das Wettbewerbsrecht als zusätzliche Herausforderung erkannt. Dabei scheint es sich jedoch um einen anderen Grund als der unlauteren Absprachen zu handeln (s.o.). Aussagen aus Interviews mit den Pilotprojekten und aus den Fachgruppenworkshops deuteten zumindest an, dass manche Pilotprojekte befürchten, dass sie durch die Vermittlung von Daten an nur einzelne Datennutzende den Wettbewerb verzerren könnten. Dem müsste eventuell noch weiter nachgegangen werden. Schließlich nennen die Pilotprojekte nun eine Reihe zusätzlicher Gesetze, die sie für einen DT als relevant ansehen. Unter diesen sticht der DGA heraus. Der Act regelt gezielt die Vermittlung von Daten. Was allerdings auffällt, sind die kritischen Stimmen vor allem aus den Experteninterviews. Der erste Kritikpunkt hier ist, dass der DGA es den DT noch schwieriger macht, den Datenteilenden bei der Bewältigung ihrer rechtlichen Herausforderungen zu helfen. Denn DT müssen nicht nur – wie die Datenteilenden – die Vorschriften zum Datenschutz, zur IT-Sicherheit und die Geschäftsgeheimnisse einhalten, sondern außerdem die Vorschriften des DGA (sowie weiterer Gesetze). Kritisiert wird dabei auch, dass der DGA den DT verbietet, die vermittelten Daten auch für eigene Zwecke zu verwenden. Zum einen fragt sich damit, wieso der DGA gerade solche Konstellationen regelt, die im Datenschutzrecht als Auftragsverhältnis erfasst sind und dabei in der Praxis kaum zu

Problemen führen. Problematisch sind dagegen solche Verarbeitungsverhältnisse, wo ein Vermittler der Daten, diese auch für eigene Zwecke verwendet. Gerade diese Konstellationen sind aber vom Anwendungsbereich des DGA ausgeschlossen. Kritisiert wird in diesem Zusammenhang auch der ökonomisch verengte Spielraum, der den DT bleibt. Dies ist in Hinsicht auf den hohen Aufwand (s.o.) und die Refinanzierung der DT-Dienste ein kritischer Faktor für den Erfolg der Dienste (siehe dazu im Detail in QT 3 und 4). Bei manchen Pilotprojekten führt die Rigidität des DGA dazu, dass diese versuchen, nicht als Vermittler von Daten zu gelten und so nicht in den Anwendungsbereich des DGA zu geraten.

In diesem Zusammenhang sind noch die Ergebnisse auf die letzte Frage aus dem rechtlichen Teil der Online-Studie interessant. Hier wurden die Pilotprojekte gefragt, mit welchen Ansätzen sie versuchen, den Aufwand bzw. die Kosten für sich selbst gering zu halten, die durch Einhaltung der rechtlichen Anforderungen entstehen. Auffällig dabei ist, dass die Standardisierung für mehr Pilotprojekte eine relevante Rolle spielt als die Übernahme der Kosten durch Datenhaltende oder Datennutzende. Dies kann allerdings auch darin begründet liegen, dass sich die meisten Pilotprojekte erst jetzt zunehmend mit der Frage zur Refinanzierung ihrer DT-Dienste befassen.

Abbildung 9 Mechanismen des DT zur Reduzierung des eigenen Aufwands bzw. der eigenen Kosten



Quelle: Eigene Darstellung (Law & Innovation) auf Grundlage der Online-Befragung der Pilotprojekte

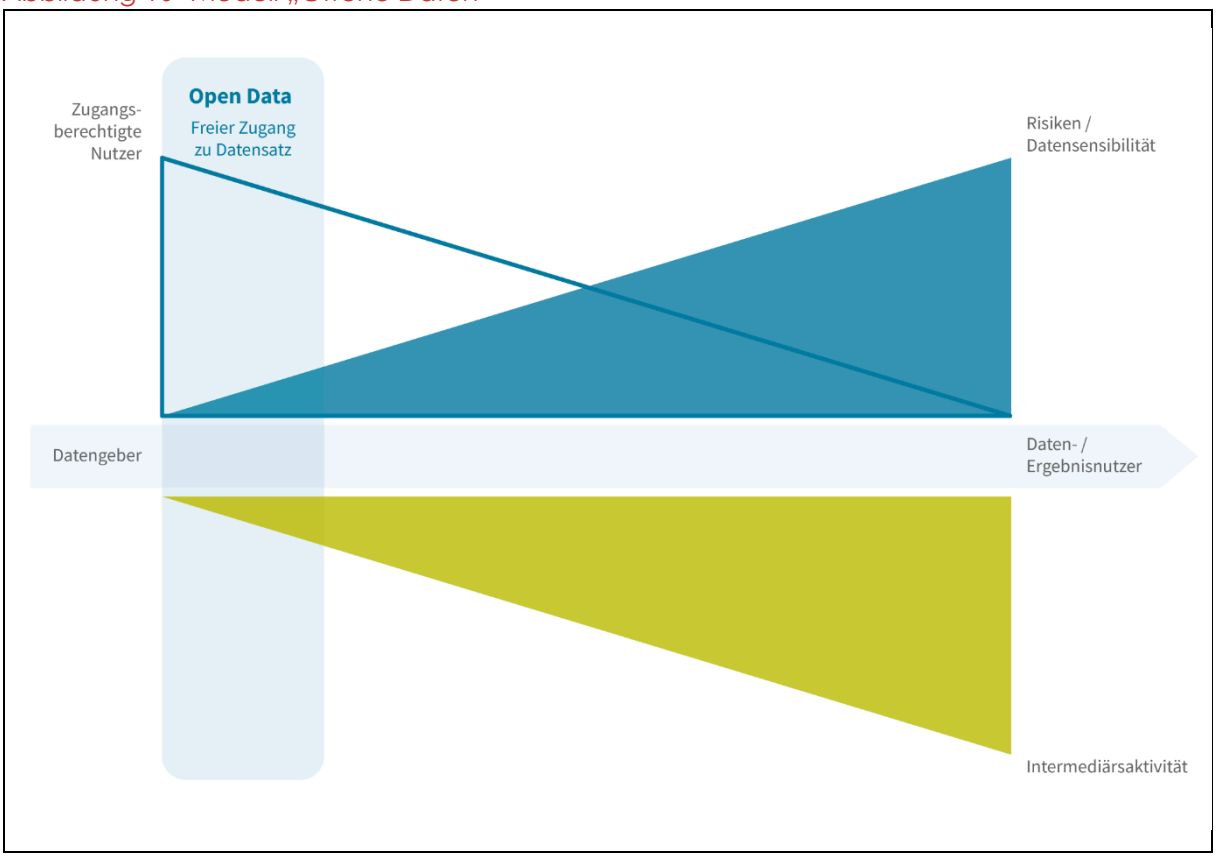
6.3 Ausblick auf konkrete DTM

Als eines der spannendsten Ergebnisse des AP1.2 wird hier angesehen, dass sich die in AP1.1 in Aussicht gestellten **drei grundsätzlichen DTM** aus rechtlicher Perspektive nicht nur bestätigt

haben, sondern sich wohl weiter systematisch ausdifferenzieren lassen. Hier wurde in AP1.1 aufbauend auf der Analyse der externen Use Cases eine erste Unterteilung in die drei folgenden Fallgruppen vorgeschlagen. Die Modelle unterscheiden sich dabei im Ausmaß der Zugangsberechtigung des Nutzenden auf die Daten bzw. der Nutzungskontrolle. Der graue Pfeil beschreibt den Vorgang des Datenteilens zwischen Datenhaltenden und Datennutzenden.

Die Zugangsberechtigung ist am umfangreichsten, wo die Compliance-Risiken des Datenhaltenden (rote Fläche) am niedrigsten sind, vgl. nachfolgendes Modell „Offene Daten" („Open Data"). Einer Intermediärstätigkeit des DT (blaue Fläche) bedarf es hier nur gegebenenfalls bezüglich organisatorischer Elemente, dagegen weniger bis gar nicht in der Stellung als Sicherheitsgarant, vorausgesetzt die Daten sind schon anonymisiert bzw. an ihnen bestehen keine anderweitigen Schutzrechte.

Abbildung 10 Modell „Offene Daten"



Quelle: Eigene Darstellung (Law & Innovation)

In den Use Cases, die dem folgenden Modell „Geteilte Daten" („Intermediär als Zugangsschranke") zuzuordnen sind, steigen die (Compliance) Risiken des Datenhaltenden an. Hierbei kann es sich um Daten handeln, denen durch rechtliche Einordnung besonderer Schutzcharakter zukommt, wie zum Beispiel Geschäftsgeheimnisse, personenbezogene Daten, etc. Umgekehrt proportional zum Anstieg der Risiken kann typischerweise die Zugangsberechtigung des Datennutzenden eingeschränkt werden. Der Datengebende verlangt bzw. benötigt hier gesteigerte Schutzmechanismen, die gegebenenfalls durch den DT sichergestellt werden. Hierbei lassen sich je grundsätzlich nach technischer und rechtlicher

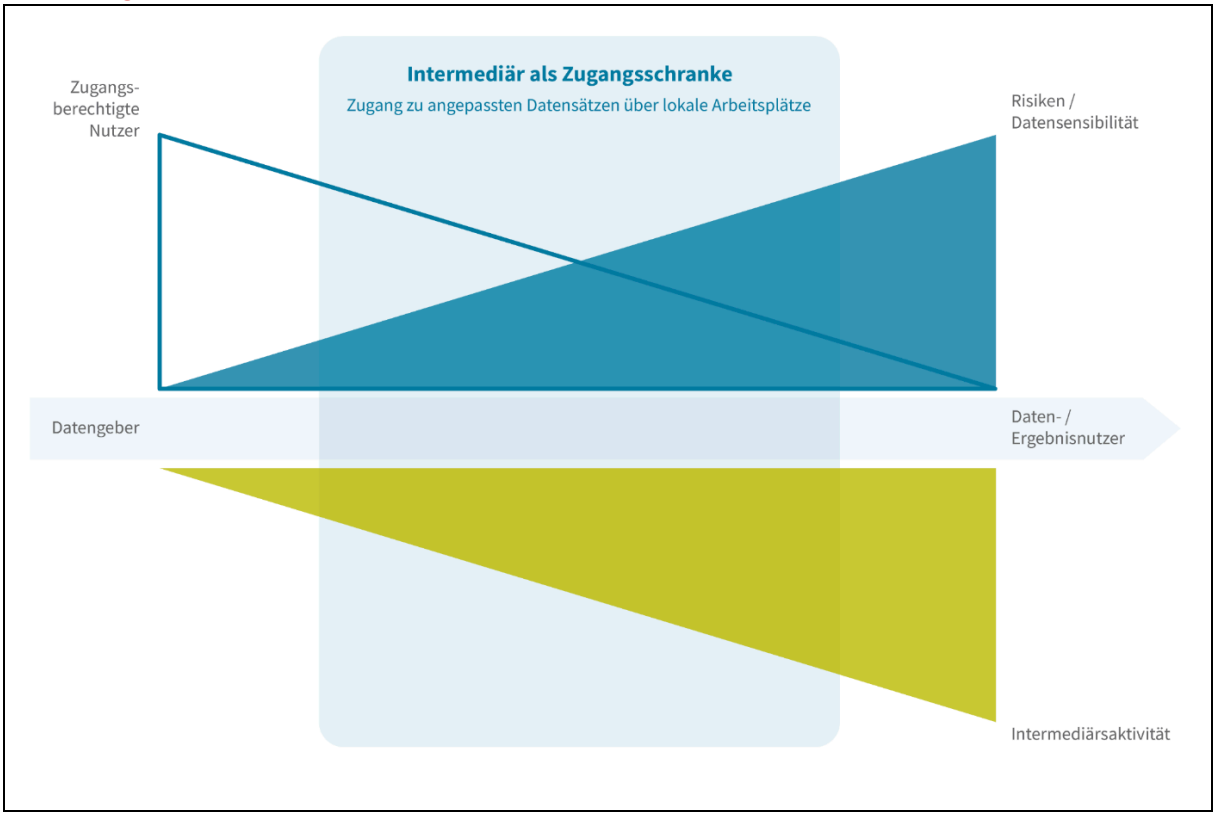
Einschränkung des Zugriffs auf die Daten verschiedene Sicherheitsstufen unterscheiden (s.a. folgende Abbildung):

Stufe 1: Formale Prüfung der Nutzungsberechtigung,

Stufe 2 (zusätzlich zu Stufe 1): Juristische Instrumente, z.B. Nutzungsvereinbarungen,

Stufe 3 (zusätzlich zu Stufe 1 und 2): Technische Instrumente, z.B. Zugang nur über lokale Arbeitsplätze.

Abbildung 11 Modell „Geteilte Daten“

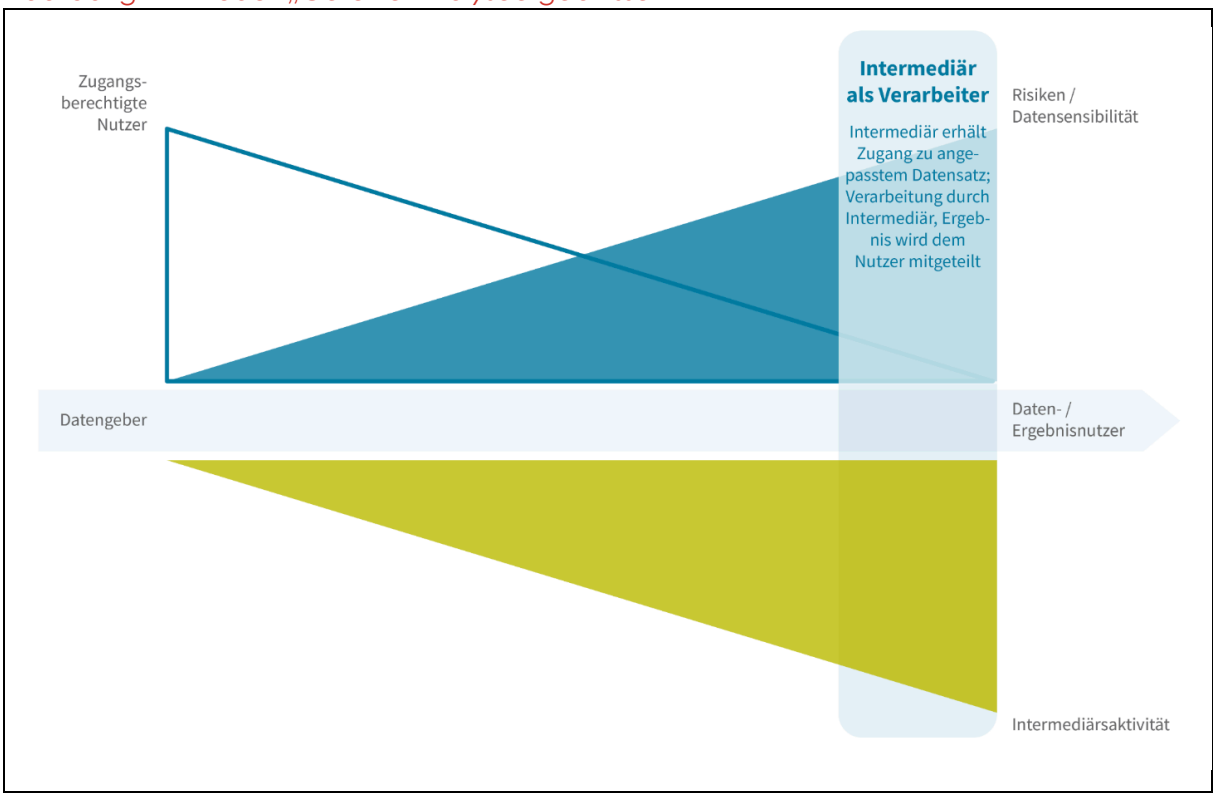


Quelle: Eigene Darstellung (Law & Innovation)

Zuletzt sieht die BF als besonderes Modell mit niedrigster Zugangsberechtigung das nachfolgende Modell „Geteilte Analyseergebnisse“ („Intermediär als Verarbeiter“). Hierbei hat der Nutzende gerade keinen Zugriff auf die Daten. Vielmehr kann er allein Anfragen an den Datenhaltenden stellen. Der DT kann dabei aus verschiedenen Gründen und auf verschiedene Weise den Datenverarbeitungsprozess leiten bzw. überwachen. Der Datenhaltende greift auf einen DT zurück, entweder weil er selbst technisch-organisatorisch nicht in der Lage ist, die Datenanalyse für den Nutzende auszuführen. Die Initiative kann aber auch vom Datennutzenden kommen, wenn dieser nicht möchte, dass der Datenhaltende die Fragen des Datennutzenden erhält. Ein solcher Fall kann etwa vorliegen, wenn der Datennutzende allein die Frage als Bestandteil seines Geschäftsgeheimnisses ansieht, das der Datenhaltende nicht erhalten soll. Wie der Grund auch sei, der DT prüft in jedem Fall das Vorliegen der zur Beantwortung der Anfrage notwendigen Daten bzw. deren Qualität. Unter Verwendung dieser Daten modelliert er ein Ergebnis passend zur Anfrage des Nutzendes (oder gibt dem Datengebenden unterstützende Anweisungen, damit dieser die Anfrage selbst beantworten kann) und überprüft lediglich das Ergebnis. Schlussendlich eröffnet er dem Nutzende das

Ergebnis des Prozesses. Der Nutzende hat in jedem Fall Zugriff nur auf das Ergebnis und wird deshalb als „ErgebnisNutzende“ eingeordnet.

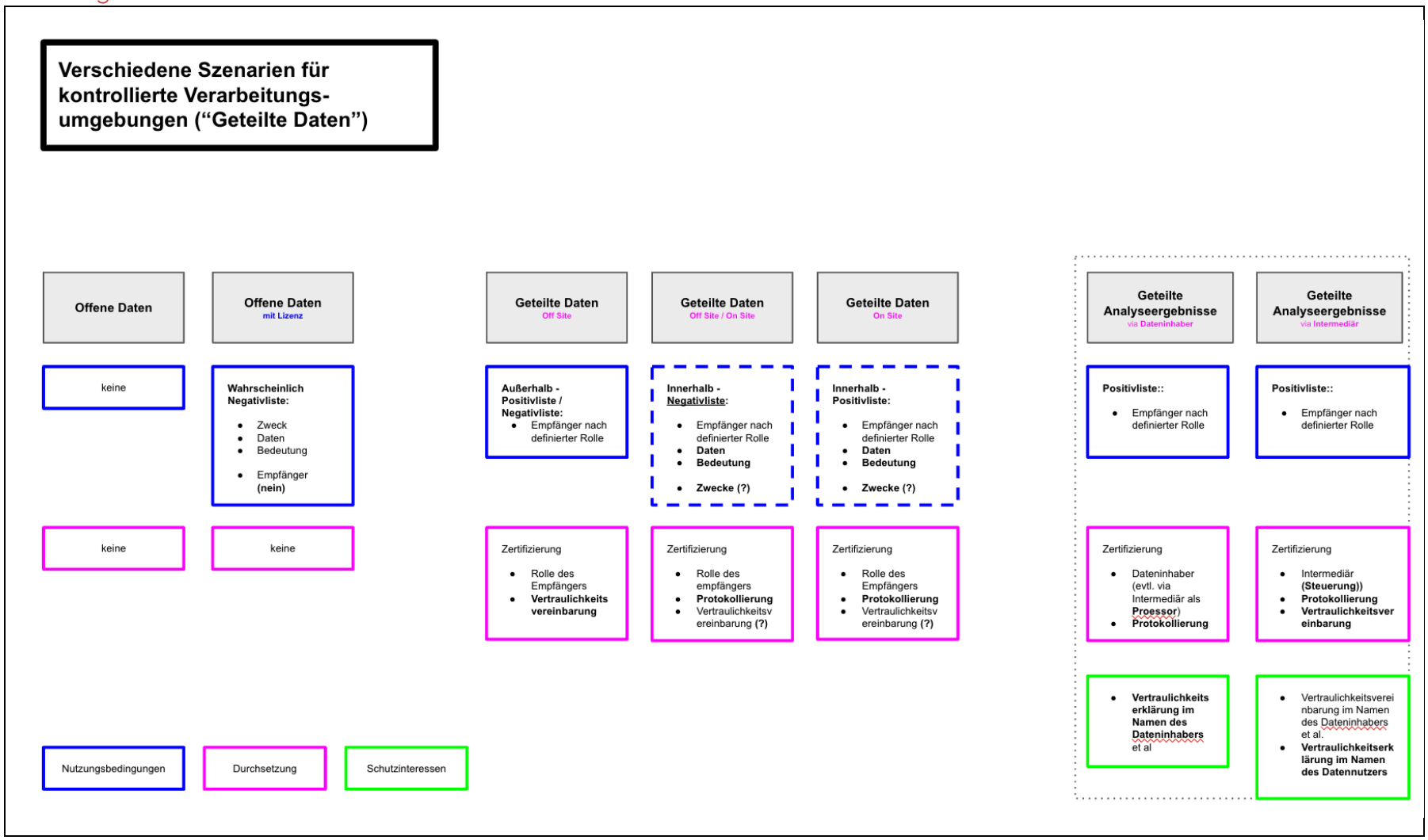
Abbildung 12 Modell „Geteilte Analyseergebnisse“



Quelle: Eigene Darstellung (Law & Innovation)

Diese erste Kategorisierung hat sich nicht nur in den Interviews mit den Pilotprojekten bestätigt. In den Fachgruppenworkshop wurde darüber hinaus ein erster Versuch unternommen, die Modelle über alle Pilotprojekte hinweg zu konsolidieren und weiter auszudifferenzieren. Dabei stellte sich heraus, dass sich die Abstufungen im Grundsatz danach richten, welche Anforderungen ein Datennutzender einhalten muss und wie deren Einhaltung überwacht wird. Anforderungen und Kontrollmechanismen sind dabei umso strenger, je größer die Risiken für den Datenhaltenden sind. So können etwa im Modell „Geteilte Daten“ nur Anforderungen an die Rolle bzw. den Status des Nutzendes gerichtet werden. In diesem Modell kann der Datennutzende die Daten in seinen eigenen Räumen mit seinen eigenen Verfahren usw. verwenden. Zusätzlich sind auch negativ oder sogar nur positiv formulierte Vorgaben für die Datennutzung möglich. Bei negativ formulierten Vorgaben dürfen Datennutzende die Daten beispielsweise nicht für bestimmte Zwecke, mit bestimmten Verfahren oder in Kombination mit bestimmten Daten verwenden; im Übrigen ist der Datennutzende aber frei. Strenger sind positiv formulierte Vorgaben, nach denen der Datennutzende die Daten nur für spezifisch erlaubte Zwecke oder mit bestimmten Verfahren oder nur in Kombination mit bestimmten Daten verwenden darf. Zumindest bei den positiv formulierten Vorgaben, darf der Nutzende die Daten meist nur in den Räumen des Datenhaltenden verwenden, um eine effektive Kontrolle zu gewährleisten. Beim Modell „Questions to the data“ werden die Rohdaten nicht geteilt, sondern nur die Ergebnisse der Analyse. Hier stellte sich eine Unterscheidung lediglich danach heraus, aus welchem Grund ein DT eingesetzt wird. Die folgende Grafik gibt einen ersten Überblick über die möglichen auszudifferenzierenden Modelle:

Abbildung 13 Wert-Risiko-Dilemma beim Teilen von Daten



Quelle: Eigene Darstellung (Law & Innovation)

Das Feedback in den Fachgruppenworkshops sowie mit dem Beirat ergab, dass eine solche Ausdifferenzierung der DTM als großen Mehrwert angesehen wird. Daher sollen diese Modelle im folgenden Begleitforschungsprozess weiter konsolidiert und ausdifferenziert werden.

7 QT3: Geschäfts- und Betriebsmodellentwicklung

Der Fokus in der folgenden Vorstellung der Befunde liegt auf der Ergänzung und „Tiefenbohrung“ in Bezug auf die in AP1.1 durchgeführte Literaturauswertung. Ferner liegt der Fokus auf Anforderungen und Herausforderungen der Förderprojekte im Bereich der Geschäftsmodellentwicklung. Die inhaltliche Schwerpunktsetzung lässt sich in die folgenden übergreifenden Themenblöcke unterteilen:

Geschäftsmodellentwicklung und DT-Angebote,

Ausgestaltung und Funktionen von Geschäftsmodellen,

Bepreisung, Zahlungsmodalitäten und Kompensation,

Übergreifende Funktionen im Datenökosystem.

7.1 Geschäftsmodellentwicklung und Datentreuhandangebote

Die Literaturanalyse in AP1.1 hat gezeigt, dass bislang eine eher gering entwickelte Landschaft an Datentreuhandanbietern existiert. Eine befragte Person (Gruppe der Expertinnen und Experten) bestätigte diesen Befund. Zwar seien Systeme bzw. Plattformen zum Datenteilen (zum Beispiel Sharing Plattformen von IBM und Amazon) bereits vorhanden, jedoch nicht ausgereift beziehungsweise ohne hohe Marktrelevanz. Die Akzeptanz dieser Plattformen unterscheide sich ferner nach Nutzendenkreis und nach Sensibilität der Daten.

Eine andere befragte Person verwies auf die hohe Anzahl an bestehenden Datenräumen und auf eine beachtliche Anzahl an DT. Aktuell wachse der Markt, würde jedoch wieder schrumpfen. Ein weiterer Interviewpartner verwies auf datenbasierte Applikationen/Services großer Anbieter, wie zum Beispiel Fitnessarmbänder, Smart Watches als Parallelstruktur bzw. als parallele Entwicklung zu DT. Auch wurden in den Interviews bestehende privatwirtschaftliche und profitorientierte Lösungen zum Datenteilen diskutiert. Während ein Interviewpartner den Eindruck äußerte, die Datenverarbeitung durch beispielsweise Amazon sei unkritisch, da Amazon ohne eigenes Domänenwissen keinen eigenen Nutzen aus Unternehmensdaten ziehen könne und die Vertragsklauseln zur Verschlüsselung ausreichend erscheinen, wies ein anderer Interviewpartner auf die großen Hemmnisse seitens Datengebenden hin, brisante Informationen mit Amazon oder IBM zu teilen, insbesondere wenn es sich um eine zentrale Datenspeicherung handele.

In der Online-Befragung gab knapp die Hälfte (sieben von 15) der Pilotprojekte an, dass ihnen Wettbewerber in ihrer Branche oder ihrem Anwendungsfeld bekannt sind, die eine **Plattform zur Vernetzung von Datengebenden und Datennutzenden** planen oder diese bereits betreiben. Wettbewerber sind vor allem in der Medizinbranche sowie der Agrar-/Forstwirtschaft bekannt. Mögliche Wettbewerber sind Initiativen des Bundes oder Bundesbehörden, wie beispielsweise das Forschungsdatenzentrum des BfArM, aber auch europäische Lösungen, wie EHDS, oder nationale Ansätze in anderen Ländern, wie FINDATA. Jedoch geben lediglich drei von 15 Befragten an, zahlreiche weitere Wettbewerber bzw. Anbieter am Markt zu erwarten.

Die Online-Befragung adressierte zudem die Hemmnisse (potenzieller) Betreiber von DT mit dem Ziel, die schwache Entwicklung der Etablierung von DTM am Markt näher zu beleuchten.

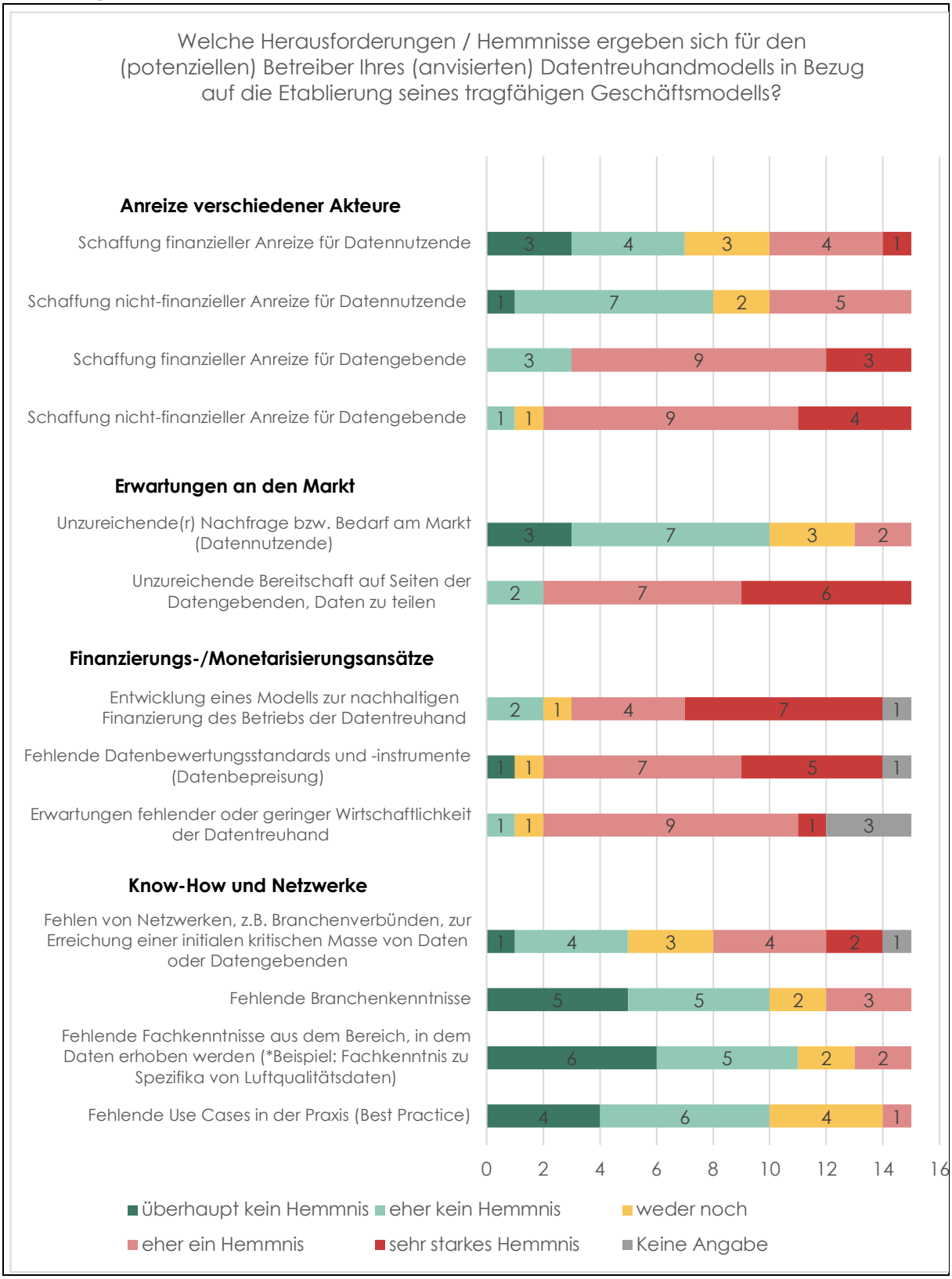
In Bezug auf Anreize verschiedener Akteure sehen die Pilotprojekte insbesondere **Herausforderungen in der Schaffung von (finanziellen wie auch nicht-finanziellen) Anreizen für Datengebenden**. Für Datennutzende hingegen scheint es weniger Bedarf zur Anreizsetzung zu geben. Diese Einschätzung spiegelt sich auch in den Erwartungen an den Markt wider. In der Nachfrage bzw. dem Bedarf potenzieller Datennutzenden werden keine signifikanten

Herausforderungen gesehen, auch hier signalisieren die Markterwartungen in erster Linie Hemmnisse in der Bereitschaft zum Datenteilen.

Laut Online-Befragung bestehen zudem Herausforderungen hinsichtlich der Monetarisierungsansätze. Hier werden vor allem fehlende Datenbewertungsstandards und -instrumente zur **Datenbepreisung** als hinderlich gesehen, aber auch die Erwartungen fehlender oder geringer Wirtschaftlichkeit der DT sowie die Entwicklung eines Modells zu deren nachhaltiger Finanzierung.

Gleichzeitig scheinen fehlendes Know-How oder fehlende Netzwerke keine besonderen Hemmnisse für DT-Betreiber darzustellen. Hierbei ist jedoch anzumerken, dass diese Aspekte im Kontext individueller Interviews durchaus als Herausforderungen genannt wurden und somit einzelfallspezifisch durchaus relevant sein können.

Abbildung 14 Hemmnisse



Quelle: Eigene Darstellung (Technopolis Group) auf Grundlage der Online-Befragung der Pilotprojekte (n=15)

Ein Interviewpartner merkte zudem an, dass viele potenzielle Anbieter Daten nicht „anfassen“ wollten, aus Angst, dass Daten geleaked werden und ein Reputationsschaden folge. Hier bestehe ein Hemmnis in der Etablierung von DT-Angeboten. Eine befragte Person (Gruppe der Experten und Expertinnen) unterstrich, dass die Nähe zu Datengebenden und Datennutzenden wesentlich sei.

7.2 Ausgestaltung und Funktionen von Geschäftsmodellen

Die Ausgestaltung von Geschäftsmodellen ist zum einen abhängig von den Vorgaben des DGA, zum anderen von den Anforderungen seitens der Marktteilnehmenden. Letzteres umfasst insbesondere die Schaffung von Vertrauen und Akzeptanz.

7.2.1 Geschäftsmodelle unter dem DGA

Im DGA sind Datenintermediäre (konkret „Datenvermittlungsdienste“) vorgesehen. Unter dem DGA darf ein Anbieter von Datenvermittlungsdiensten die Daten, für die er Datenvermittlungsdienste erbringt, für keine anderen Zwecke nutzen, als sie den Datennutzenden zur Verfügung zu stellen (DGA, Art. 12a).

Datenvermittlungsdienste können ein Angebot zusätzlicher spezifischer Werkzeuge und Dienste für Dateninhabende oder betroffene Personen umfassen, insbesondere um den Datenaustausch zu erleichtern, z.B. die vorübergehende Speicherung, Pflege, Konvertierung, Anonymisierung und Pseudonymisierung; diese Werkzeuge werden jedoch nur auf ausdrücklichen Antrag oder mit Zustimmung des Dateninhabenden oder der betroffenen Person verwendet, und die in diesem Zusammenhang angebotenen Werkzeuge Dritter werden für keine anderen Zwecke verwendet (DGA, Art. 12e). Datenaltruistischen Organisationen unterliegen weiteren Beschränkungen des DGA (DGA, Art. 17).

In der Literaturanalyse wird häufig die Meinung vertreten, dass darüber hinaus kein kommerzielles Interesse bestehen dürfe. Im Gesetzestext selbst findet sich dies allerdings nicht wieder. **Ein gewinnorientiertes Geschäftsmodell ist somit grundsätzlich zulässig.** Auch wird keine wirtschaftliche Unabhängigkeit verlangt. Der Anbieter des Datenvermittlungsdiensts darf jedoch seine Leistung nicht von der Nutzung anderer eigener Dienste abhängig machen.

Eine befragte Person (Gruppe der Expertinnen und Experten) merkte an, dass die verabschiedete Fassung des DGA „flacher“ geworden sei als ursprünglich intendiert. Dies zeige sich auch darin, dass die Anforderungen an die Datenvermittlungsdienste in Bezug auf Neutralität im Gesetzestext selbst deutlich schwächer ausfallen als in den vorangestellten „Erwägungen“.

In den Interviews wurden die **Anforderungen in Bezug auf Neutralität** (wie in DGA, Artikel 12a, definiert) grundsätzlich unterstützt, wie auch der Fokus auf Anonymisierung, Harmonisierung bzw. Nutzbarmachung und Qualitätssicherung zur Wahrung von Interessen der Datengebenden. Auf der anderen Seite sahen einzelne Interviewpartner auch die eigene Datenverarbeitung durch den DT als erfolgskritisch.

In den Fachgruppenworkshops zeichneten sich größere Unklarheiten in Bezug auf die Interpretation des DGA in der Praxis ab. Selbst unter Juristinnen und Juristen scheinen die Rahmenbedingungen in diesem Kontext unklar zu sein. Eine Person schilderte eine rechtlich unklare Situation in der Medizin, in der Einwilligungserklärungen mit personenbezogenen Daten (z.B. Geschlecht, Alter, Standort/Klinik, jedoch nicht verknüpft mit Krankheitsbild) vorliegen, es jedoch unklar bliebe, ob diese Daten für Statistiken verwendet werden dürfen. Ferner würden Dienstleistungen wie z.B. Datenaufbereitungsdienste, oder „match-making“ Funktionen eingeschränkt. Kern der Unklarheiten zur Auslegung des DGA liegt in der Definition der Datenanalyse für eigene Zwecke, welche der DGA beschränkt. Eine Person merkte an, dass

selbst die Datenspeicherung und -anonymisierung für Juristinnen und Juristen bereits unter datenverarbeitende Maßnahmen fallen könnten. Grundsätzlich erscheinen rechtliche Aspekte die größten Hemmnisse der Etablierung von Geschäftsmodellen. Dennoch herrschte kein Konsens darüber, ob der DGA de facto ein Hindernis darstellt. Eine Person merkte an, Konformität mit dem DGA sei möglich, indem ein DT als juristische Person gegründet wird. Dieser Aufwand könne jedoch eine Einstiegshürde darstellen.

7.2.2 Organisationsform

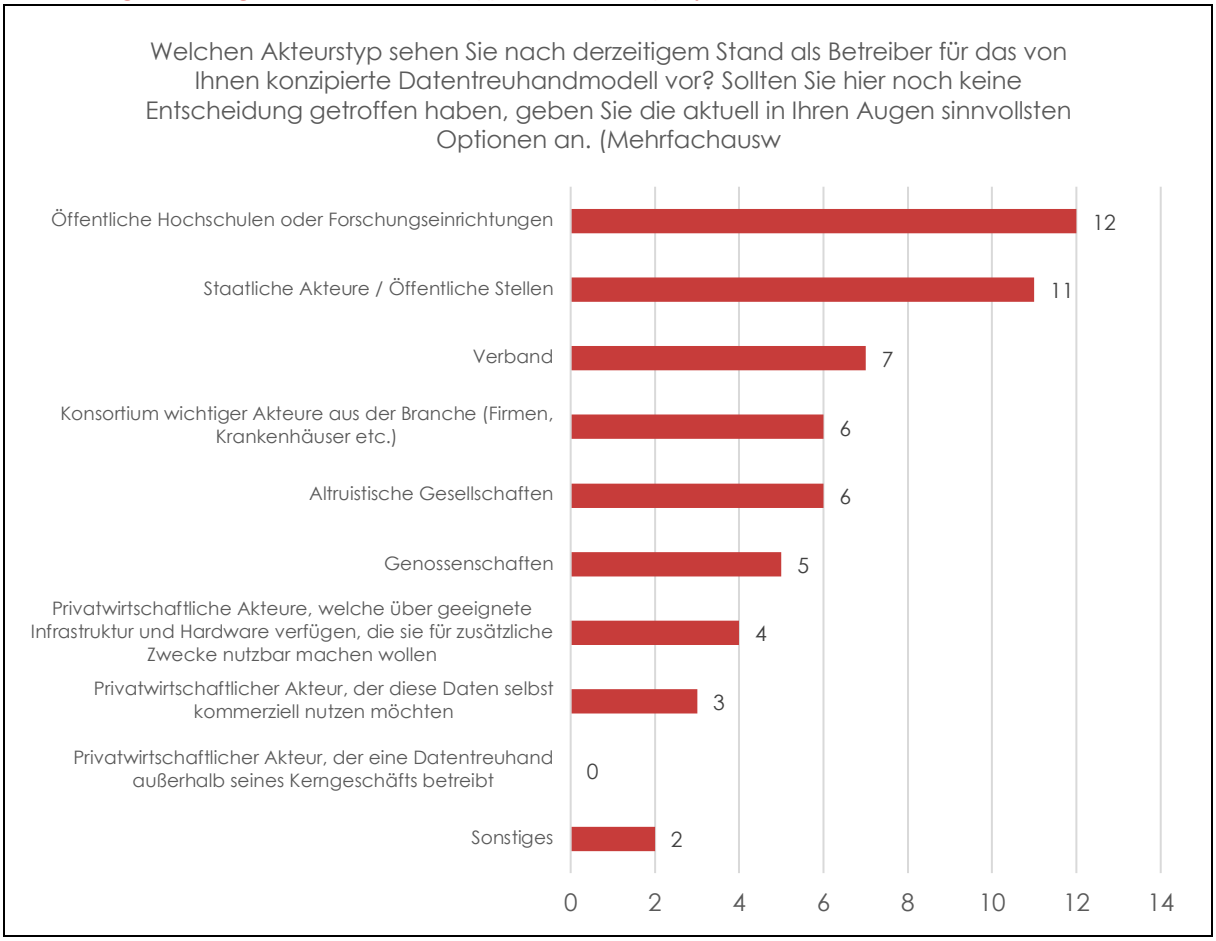
In der Literatur wird ein DT meist als unabhängige Instanz verstanden, die den Ausgleich unterschiedlicher – und oft widersprüchlicher – Anreize zur Datenbereitstellung und -verarbeitung sowie gemeinsamer Datennutzung ermöglichen soll, während die Interessen und das Recht auf informationelle Selbstbestimmung aller Stakeholder gewahrt werden.

Eine befragte Person (Gruppe der Experten und Expertinnen) unterstrich die **Funktion eines DT** als zwischengeschaltete Vertrauensstelle und „unabhängige Instanz" zwischen Datengebenden und Datennutzenden, die vor allem die Bedürfnisse und Interessen der Datengebenden wahrt, sowie gewährleistet, dass Daten rechtlich legitim verwendet werden. Letzteres solle auch im Sinne von Ethik und Governance begriffen werden, sodass in die Datennutzung auch normative Kriterien miteinfließen. Die befragte Person unterstrich insbesondere die Anforderungen an einen DT im Gesundheitswesen. Hier stünden ethische Fragen der Datenverarbeitung und des Datenzugangs (z.B. öffentliche Forschung vs. Pharma-Industrie) im Vordergrund. Eine weitere befragte Person vertrat die Ansicht, eine Treuhänder-Funktion für personenbezogene Daten müsse eine öffentliche Stelle übernehmen, bei nicht-personenbezogenen Daten seien auch andere Betreiber möglich.

In der Online-Befragung zeigt sich, dass die Mehrheit der Pilotprojekte **öffentliche Einrichtungen als Betreiber für ihre DTM** vorsehen. Wie Abbildung 15 zeigt, sehen zwölf von 18 Pilotprojekten öffentliche Hochschulen oder Forschungseinrichtung, elf von 18 staatliche Akteure oder öffentliche Stellen vor. Die befragten Pilotprojekte gaben zudem mehrheitlich an, dass es für die die Akzeptanz des Betreibers wichtig ist, dass ein öffentlicher/staatlicher Akteur, den DT betreibt. Dieser Aussage stimmten zwölf von 16 Pilotprojekten zu. Privatwirtschaftliche Akteure werden nur von wenigen Pilotprojekten als Betreiber vorgesehen.

In der Online-Befragung gaben nur wenige Pilotprojekte (fünf von 18) an, dass sie eine Genossenschaft als Betreiber vorsehen (siehe Abbildung 15). In den Fachgruppenworkshops wurde das genossenschaftliche Betriebsmodell jedoch grundsätzlich als sehr positiv hervorgehoben. In diesem Modell entscheiden Mitglieder über die Datennutzung und partizipieren an Gewinnen.

Abbildung 15 Vorgesehene Akteurstypen der Pilotprojekte



Quelle: Eigene Darstellung (Technopolis Group) auf Grundlage der Online-Befragung der Pilotprojekte (n=18)

In den Fachgruppenworkshops wurden ebenfalls die **Anforderungen an die Organisationsform und die Ausgestaltung zur Gewährleistung von Neutralität und Akzeptanz in der Praxis** diskutiert. Auf der einen Seite wurde hierbei unterstrichen, dass sich der Stand der Technik (zum Beispiel in Bezug auf Anonymisierung) kontinuierlich ändere und private Akteure diesbezüglich innovativer seien als staatliche. Auf der anderen Seite wurde insbesondere im medizinischen Bereich auf Probleme der Akzeptanz und des Vertrauens gegenüber einem privaten DT verwiesen. Hier zeigt sich in der gegenwärtigen Diskussion ein „Trade-Off" zwischen den eher als vertrauensvoll angesehenen öffentlichen Betreibenden gegenüber der höheren technischen Kompetenzvermutung seitens privater Anbieter.

Ferner wurde in den Fachgruppenworkshops diskutiert, ob die relevante Unterscheidung in Bezug auf die Akzeptanz nicht „privat vs. staatlich", sondern vielmehr **„for-profit versus not-for-profit"** sei. Es herrschte Uneinigkeit darüber, ob ein profitorientiertes Modell Akzeptanz grundsätzlich gewährleisten kann. Als Argument für ein profitorientiertes Modell wurde die Möglichkeit der Re-Investition genannt. Außerdem wurde von einem Teilnehmenden argumentiert, dass nur ein kommerziell orientierter DT die Mittel habe, die technisch aufwendige Anonymisierung von Daten zu gewährleisten. Zudem erfolge die Entscheidung der Datengebenden nach dem Verhältnis von Kosten (auch nicht monetär, zum Beispiel Reputationsverlust) und Nutzen. In der Frage der Akzeptanz sei das entscheidende Kriterium somit der Mehrwert, die Finanzierung sei demnach eine Frage von Angebot und Nachfrage.

Auch im Medizinbereich gäbe es beispielsweise in den USA bereits weniger Hemmungen, Patienten-Daten zu monetarisieren. In Europa und Deutschland sehe man dies jedoch grundsätzlich kritischer. Es wurde auch die Frage aufgeworfen, ob es entscheidend ist, mit welchen Funktionen ein DT Profite erzielt – mit den Daten selbst oder lediglich über die Transaktion dieser. Letzteres sei aus Sicht der Datengebenden und -nutzenden weniger problematisch. Es wurde außerdem das Modell einer Genossenschaft für den DT vorgeschlagen, in der die Profite an die Teilhaber zurückfließen würden.

Andere Teilnehmende sprachen sich insbesondere im Medizinbereich für ein nicht-profitorientiertes Modell aus. Auch die geführten Interviews deuten an, dass konkret im Gesundheitsbereich ein staatlicher Akteur sowie eine nicht-kommerzielle, nicht-profitorientierte Lösung vorzuziehen sei. Gründe umfassen sowohl ethische Bedenken wie auch Akzeptanz. Ferner wurde im Rahmen der Fachgruppenworkshops unterstrichen, dass es im Medizinbereich in Deutschland bislang an einer entsprechenden ethischen Diskussion mangele. Grundsätzlich wurde vorgeschlagen, in besonders sensiblen Bereichen, wie der Medizin, eine dritte Instanz oder eine Ethikkommission einzurichten, die bei rechtlichen/umstrittenen Fragen vermittelt.

Die Online-Befragung ergab, dass befragte Pilotprojekte es mehrheitlich (zwölf von 16) eher wichtig für die Akzeptanz des DT erachten, dass dieser auf non-profit-Basis betrieben wird. Die überwiegende Mehrheit der Pilotprojekte (zehn von 16) sehen zudem laut Online-Befragung keine kommerzielle Orientierung für ihren (anvisierten) DT vor.

Für das **Betriebsmodell** zeigt die Befragung, dass der Großteil (13 von 16) es als wichtig erachtet, dass der DT neben der Datenaufbereitung und Vermittlung des Datentauschs gleichzeitig Dienste der Anonymisierung/Pseudonymisierung der Daten anbietet. Auch wird es mehrheitlich als wichtig für die Akzeptanz angesehen, dass keine Abhängigkeiten oder direkten Anbindungen zur Wertschöpfungskette des DT-Betreibers bestehen und eine sichtbare Trennung der Treuhänder-Funktion vom Kerngeschäft des Betreibers besteht (jeweils neun von 16 Pilotprojekten stimmen dieser Aussage zu). Die Pilotprojekte stimmen eher zu, dass es die Akzeptanz und Nutzung des DT bei Datennutzenden steigern würde, wenn der DT neben der Datenaufbereitung und Vermittlung des Datenaustauschs noch weitere zusätzliche Dienste anbieten würde, wie zum Beispiel Datenauswertungen oder Bereitstellung von Analytics-Tools. Gleichzeitig sehen sie auch wenig Risiko, dass die Akzeptanz bei Datengebenden durch solche Angebote sinken würde.

Unabhängig vom anvisierten Betreibermodell plädiert die Mehrheit für staatliche Anschubfinanzierung zur Etablierung eines nachhaltigen Geschäftsmodells, wie 14 von 16 Pilotprojekten angaben.

Zusammenfassend zeigt sich, dass die Geschäftsmodellausgestaltung je nach Branche unterschiedliche Anforderungen aufweisen muss. In der Befragung stimmten 14 von 15 Pilotprojekten dieser Aussage zu.

Dieser Befund wurde in den Interviews mit Expertinnen und Experten bestätigt. Zudem sei eine Pluralisierung von DTM mit unterschiedlicher Zugangsoffenheit wichtig, da manche Datengebende zu einer höheren Freigiebigkeit ihrer Daten bereit seien als andere. Laut einer befragten Person (Gruppe der Expertinnen und Experten) sei noch unklar, ob sich der Markt in Richtung domänenbezogenen oder übergreifenden DT entwickeln werde, aktuell gehe der Trend hier auseinander.

Hinsichtlich der **langfristigen Finanzierung nach Förderende** sehen die meisten Pilotprojekte (14 von 18) vor allem eine Finanzierung durch Datennutzende, z.B. durch Gebühren für den Zugriff auf Daten oder für Dienstleistungen, sowie zehn von 18 durch öffentliche Finanzierung. Weniger

relevant erscheinen Zuwendungen von Interessensgruppen oder von nicht-staatlichen Organisationen (fünf beziehungsweise 3 von 18).

7.3 Bepreisung, Zahlungsmodalitäten und Kompensation

Bei der **finanziellen Ausgestaltung von Geschäftsmodellen** scheint es in der Praxis wenige Erfolgsbeispiele zu geben. Die vom BMBF geförderten Projekte befinden sich mehrheitlich bislang noch nicht im Stadium eines ausgereiften Konzepts. Auch die Aussagen in der Literatur zeigen sich bislang relativ „oberflächlich" in ihren Empfehlungen.

Die **Bepreisung von Daten** stellt der Literatur zufolge insbesondere im C2B-Kontext noch eine große Herausforderung dar. Es mangle demnach an Datenbewertungsstandards und -instrumenten. Grundsätzlich wird in der Literatur empfohlen, Geschäftsmodelle nicht auf die Bepreisung der Daten selbst zu stützen, sondern auf Gebühren. Die Wahl eines geeigneten Finanzierungsmodells hänge stark von Faktoren wie dem Zweck des DT, dem Typ und der Anreize von Datenbereitstellenden, sowie der Arten der Nutzung und der Nutzenden ab.

7.3.1 Zahlungsmodalitäten

In der Online-Befragung zeigt sich, dass die Pilotprojekte überwiegend ein **Subskriptionsmodell als Preismodell** für ihren DT sehen (siehe Abbildung 16). Dies finden fünf Pilotprojekte sehr und sieben Projekte eher zielführend. Auch eine **Bepreisung abhängig vom geteilten Datenvolumen** erachten 12 von 16 Projekten als sehr beziehungsweise eher zielführend. Am wenigsten relevant sind für die Befragten Modelle, in denen eine Gebühr oder Transaktionsgebühr pro Nutzung oder Zugriff erhoben werden. Jeweils 9 von 16 empfinden diese sehr beziehungsweise eher zielführend. Mehrheitlich sehen die Befragten (10 von 16) zielgruppenspezifische Preismodelle für Datennutzende als sinnvoll an. Mehrheitlich sehen die Projekte eine Differenzierung nach Wissenschaft, Wirtschaft und öffentlichen Akteuren. Preisreduktionen oder eine kostenfreie Nutzung von Akteuren aus dem öffentlichen Sektor oder der Forschung werden vereinzelt angedacht.

Eine befragte Person (Gruppe der Experten und Expertinnen) sagte aus, sie beobachte eine Marktentwicklung in Richtung hin zu Flatrate-Modellen.

Bei der **Preissetzung** sehen die Projekte vor allem **kostenbasierte Ansätze** zielführend, das heißt der Preis wird auf der Grundlage der Kosten und einer zusätzlichen Marge festgelegt, dies gaben 13 von 16 Projekten an. Acht Projekte sehen eher einen wertorientierten Ansatz zielführend und vier Projekte eine Marktorientierung.

Ein Interviewpartner zweifelte die Zahlungsbereitschaft der Industrie für DT an. Eine befragte Person (Gruppe der Expertinnen und Experten) wies auf die Schwierigkeit von zusätzlichen Kosten durch „Zwischenschalten" einer Datentreuhand hin. So bestimmten Datennutzende (insbesondere „Big Tech") lieber selbst über die Datennutzung, eine Abhängigkeit von einer zusätzlichen Instanz würde auf Ablehnung stoßen. Dies erschwere die Etablierung eines Geschäftsmodells für eine Datentreuhand. Ferner verwies sie auf das „Capture by the client"-Phänomen, also das Risiko bei finanzieller Abhängigkeit von Datennutzenden. Hier sei die Anpassung an die Bedürfnisse der Datennutzenden ein natürliches Phänomen.

7.3.2 Kompensation für Datengebende

In der Literatur findet sich ein breiter Konsens darüber, dass von monetären Anreizen bzw. Vergütung von datengebenden Individuen (B2C Kontext) abzusehen sei, da durch eine solche Vergütung konträre soziale Effekte zu erwarten sind. Ferner sei eine solche **Datenbepreisung** nicht sinnvoll, da die Daten eines einzelnen Individuums ohnehin keinen großen Mehrwert hätten.

Ein Interviewpartner erklärte, es sei schwierig, einen Preis für die monetäre Vergütung von Datengebenden zu finden, der akzeptiert wird, da der Nutzen des Datennutzenden ein Geschäftsgeheimnis und damit nicht transparent ist.

Bei der Vergütung bzw. Kompensation für Datengebende sehen die befragten Pilotprojekte vorrangig eine monetäre Vergütung für die bereitgestellten Daten (10 von 16) und/oder die Erstattung des Aufwands bzw. der Kosten für die Bereitstellung der Daten (10 von 16). Eine altruistische Datenspende, also keine monetäre oder nicht-monetäre Vergütung sehen lediglich 5 Projekte vor. Eine Kompensation über eine unentgeltliche oder kostenreduzierte Nutzung von Diensten des DT oder über den Zugriff auf bzw. Zugang zu Daten anderer Datengebenden sehen auch lediglich sechs bzw. fünf Projekte vor.

Ein Interviewpartner merkte im B2B-Kontext an, monetäre Anreize funktionierten besser. Ein reiner Datentausch als Konzept sei schwieriger, da Daten nicht immer gleichermaßen benötigt würden.

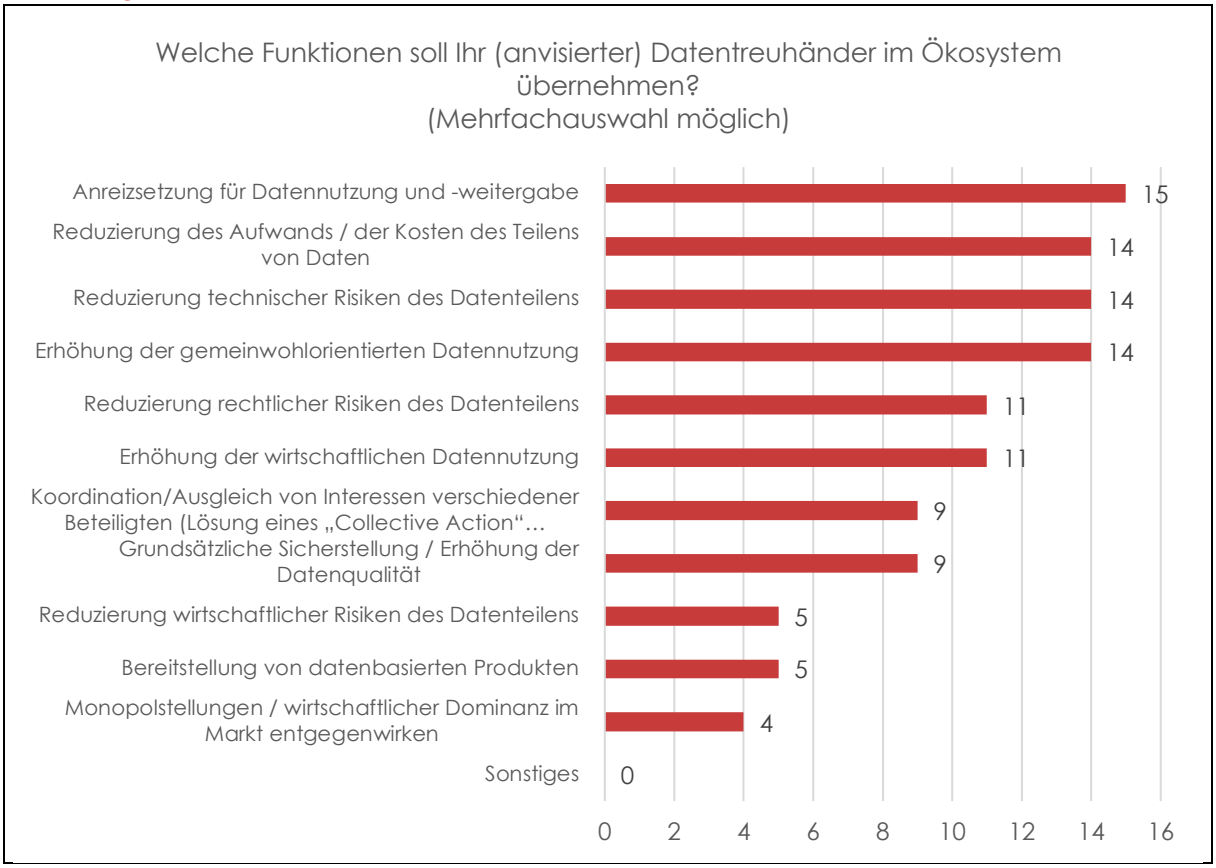
Ein Interviewpartner schlug einen Community-basierten Ansatz zur Bepreisung von Daten vor, wobei Mitglieder der verteilten Infrastruktur kooperativ den Wert der zur Verfügung gestellten Daten bewerten.

7.4 Übergreifende Funktionen im Datenökosystem

Die Literaturauswertung zeigt, dass DT im Datenökosystem eine zentrale Funktion einnehmen können, indem sie Datennutzende und Datengebende zusammenführen.

Die Pilotprojekte sehen die **Funktion** vor allem in der Anreizsetzung für Datennutzung und -weitegaben, Erhöhung der gemeinwohlorientierten Datennutzung, Reduzierung technischer Risiken und des Aufwands des Datenteilens (siehe Abbildung). Weniger relevant sind laut Online-Befragung die Bereitstellung von datenbasierten Produkten, die Reduzierung wirtschaftlicher Risiken des Datenteilens und das Gegenwirken von Monopolstellungen oder wirtschaftlicher Dominanz am Markt.

Abbildung 16 Funktionen des Datentreuhänders im Datenökosystem



Quelle: Eigene Darstellung (Technopolis Group) auf Grundlage der Online-Befragung der Pilotprojekte (n=16)

In den Fachgruppenworkshops wurde als zentrale Aufgabe eines DT im Ökosystem vor allem die **technische und organisatorische Förderung des Datenaustauschs und die Einnahme einer Vermittlerrolle** unterstrichen. Datenmissbrauch sei das Hemmnis, Daten zu teilen, dieses zu beheben sei die zentrale Aufgabe eines DT. Dennoch solle ein DT im Auftrag beider Seiten handeln und als Vermittler sowohl Aufgaben im Sinne der Datengebenden als auch der Datennutzenden übernehmen. Auf der Seite der Datengebenden solle ein DT die Freigabe der Daten kontrollieren, sowie die Nachverfolgbarkeit gewährleisten und kontrollieren. Auf der Seite der Datennutzenden solle ein DT vor allem die Verfügbarkeit und Nutzbarmachung von Daten übernehmen, Kriterien für Datensätze festlegen, Metadaten anbieten, Datenquellen verifizieren und gegenbenenfalls Suchmasken erstellen.

In der Übernahme genannter Aufgaben bestünden laut Aussagen im Kontext der Fachgruppenworkshops jedoch diverse Herausforderungen. So bestehe ein gewisses Spannungsverhältnis zwischen den Anforderungen einer vollständigen Anonymisierung/ Pseudonymisierung auf der einen Seite sowie dem rechtlichen Anspruch auf die spätere vollständige Löschung von Daten seitens Datengebenden. Im Falle eines Löschungsanspruchs müssten so ganze Datensätze unter Ausschluss der Daten des Einzelnen vollständig neu erstellt werden. Entsprechend bedürfe es einer Verknüpfung über alle Geschäftsprozesse und Datenerhaltenden hinweg, um den Ausschluss der Datennutzung ab Zeitpunkt des Widerrufs zu gewährleisten. Zum anderen bestünde eine Herausforderung im bedarfsorientierten Auffinden von Daten für Datennutzende auf Basis von nicht einheitlichen Datenbeschreibungen (z.B. Multilingualität). Standardmasken seien oftmals nicht ausreichend

feingliedrig. Die Strukturierung dieses komplexen Suchprozesses solle Aufgabe des DT sein. Entwicklungen im Bereich der Künstlichen Intelligenz deuten darauf hin, dass diese Suchprozesse möglicherweise langfristig von verbesserten Algorithmen übernommen werden könnten.

Ergänzend wurde auf den Aufwand durch die Neustrukturierung der Daten vor Weitergabe verwiesen. Durch Privacy-by-Design Anforderungen würden Daten anonymisiert gespeichert, sobald Daten personenbezogen weitergegeben werden (im Sinne der Nutzbarmachung), müsse die Struktur erst einmal gewandelt werden. So werde in der Automobilbranche beispielsweise argumentiert, dass Daten in der vorliegenden Struktur nicht nutzbar für eine Weitergabe seien. Der hohe Aufwand der Datenstrukturierung müsse entsprechend einem monetären Gegenwert gegenüberstehen.

In der Online-Befragung zeigt sich, dass die Pilotprojekte eine Vielzahl von **Leistungen** planen und auch jeweils mehrere Leistungen anbieten (werden). Fast alle befragten Pilotprojekte bieten die Verwaltung von Datenzugriffsrechten und Nutzungskontrollen (15 von 16), die Bereitstellung einer IT-Infrastruktur oder technischer Bausteine zur Datenübertragung und Anonymisierung, Pseudonymisierung und/oder Verschlüsselung an (jeweils 13 von 16). Weniger häufig werden die zentrale Datenspeicherung bzw. Hosting, die Harmonisierung und die Datenaufbereitung von Rohdaten angeboten (jeweils 7 von 16).

Zum **Erreichen einer kritischen Masse** sei es grundsätzlich ausschlaggebend, dass die richtigen Akteure im Ökosystem teilnehmen. Eine Herausforderung läge darin, Barrieren zu überwinden, indem genug Anreize zur Teilnahme gesetzt werden. Eine kritische Masse definiert sich laut einem Teilnehmenden in der Möglichkeit, Rückverfolgbarkeit auszuschließen. Auch Verbände könnten eine Rolle in domänenspezifischen Ökosystemen einnehmen. In der Online-Befragung wird das Fehlen von Netzwerken, zum Beispiel Branchenverbünden, zur Erreichung einer initialen kritischen Masse von Daten oder Datengebenden als eher unkritisch angesehen.

7.4.1 Monopole

Die Literaturauswertung deutete darauf hin, dass neue Governance-Modelle im Bereich DT Machtungleichgewichte am Markt teilweise beseitigen und einen offeneren Datenzugang unter Datenschutzvoraussetzungen wie auch die Entwicklung eines fairen Markts fördern können. Dafür müsse vermieden werden, durch Lock-in Effekte oder Datensilos eine monopolähnliche „Supertreuhand" entstehen zu lassen.

Eine befragte Person (Gruppe der Expertinnen und Experten) wies darauf hin, dass **Netzwerkeffekte** einen natürlichen Trend zur Monopolisierung bewirkten. Daher könne es sinnvoll sein, durch Regulierung oder Anreize, Monopolisierungstendenzen zu brechen, wie es beispielsweise der Digital Market Act oder der Digital Service Act der EU bereits versuche. Dies könne es auch für DT geben. Das Entstehen eines branchenübergreifenden Monopols wurde von mehreren Interviewpartnern angezweifelt. Ferner wurde angemerkt, es sei zu regulieren, welche Daten aus einem öffentlichen Interesse heraus geteilt werden sollen. Ein Interviewpartner nahm dagegen die Position ein, Unternehmen mit einem sehr hohen Marktanteil (wie Google) seien nicht notwendigerweise problematisch, diese bieten aufgrund ihrer Position gute Serviceangebote. Bei Gesundheitsdaten wurde allerdings auch hier die Rolle des Staats als Instanz ohne Eigeninteresse unterstrichen.

Eine befragte Person (Gruppe der Expertinnen und Experten) merkte an, es sei unklar, ob sich der Markt in Richtung Monopolisierung entwickelt. Die Entstehung eines übergreifenden Monopols sehe sie jedoch eher nicht. Zudem wäre ein Domänenmonopol ohne eigene Verwertungsabsicht weniger problematisch, denn so würde lediglich eine rechtliche Kontrollinstanz geschaffen.

Die Teilnehmenden der Fachgruppenworkshops beobachten zurzeit keine Monopolisierungstendenzen, eher ein Netzwerk mehrerer DT und Ökosysteme, welches es zu verknüpfen gelte. Ein sektorübergreifender DT sei den Teilnehmenden zufolge weniger wahrscheinlich, es brauche sektorale Kenntnisse und Vertrauen, so sei vielmehr eine oligopolistische Struktur zwischen den Domänen zu erwarten. Eine Vernetzung von domänenspezifischen DT über Konnektoren sei aber denkbar. Ein branchenspezifisches Monopol eines DT erschien den Teilnehmenden nicht unmittelbar problematisch, zentral seien vor allem ausreichende Kontrollmechanismen. Letztere könnten durch genossenschaftliche Organisationsstrukturen ermöglicht werden, in denen Mitglieder über die Datennutzung entscheiden.

Einige Workshop-Teilnehmende plädierten zudem dafür, das Thema DT solle sich zunächst einmal entwickeln und in den breiten Markt ankommen, bevor Debatten über Monopolisierung geführt werden sollten. Andere verwiesen auf die späte Reaktion auf die Monopolisierung der Plattformökonomie. Grundsätzlich habe ein Ökosystem mit starken Netzwerkeffekten einen natürlichen Hang zur Monopolisierung. Hier bestehe auch, wie oben bereits dargelegt, immer ein Trade-off zwischen Innovationskraft und Regulierung, als Herausforderung für staatliche Intervention.

7.4.2 Anreize und Hemmnisse für Datengebende und Datennutzende

Interviews zum Thema Geschäftsmodelle behandelten auch das Thema **Angebot und Nachfrage im Ökosystem**.

Während die Partizipation von Datengebenden erfolgskritisch sei, zählen zu möglichen **Hemmnissen** zum Beispiel die hohe Sensibilität von Daten (insbesondere Gesundheit, IT-Sicherheit), ein potenzieller Reputationsverlust, potenzielle Daten-Leaks, ein hoher Aufwand der Datenbereitstellung, potenzielle Wettbewerbsvorteile der Konkurrenz, sowie Informationsasymmetrien über andere Beteiligte im Datenpool. In Bezug auf letzteres wurde auf ein potenzielles Dilemma verwiesen. Zum einen sei es wesentlich, dass die „großen Player" im Datenpool vertreten sind, sodass eine kritische Masse beziehungsweise ein Mehrwert gegeben sind, andererseits dürfen keine Rückschlüsse zur Herkunft aus den Daten gezogen werden. Bei branchenspezifischen Pools sei dies potenziell schwierig.

Auf Seiten der Datennutzenden zeichnen sich zwar diverse potenzielle Use Cases ab, dennoch wurde mehrfach angemerkt, der Nutzen einer DT seitens Datengebenden sei nicht immer direkt ersichtlich. Hier seien Best Practice Beispiele nötig, um Unternehmen die konkreten Mehrwerte aufzuzeigen sowie das Verhältnis zum Aufwand, und zur Überwindung des „Henne-Ei-Problems" darzulegen. Fehlende Use-Cases werden in der Online-Befragung eher nicht als Hemmnis angesehen, so sehen 10 von 15 Projekten hier kein Hemmnis.

Als Hemmnisse seitens der Datennutzenden wurden in den Interviews insbesondere hohe Anforderungen an Verlässlichkeit und Qualität der Daten genannt, die gegebenenfalls strikte Qualitätsstandards und -kontrolle voraussetzen, insbesondere im Bereich IT-Sicherheit, medizinische Daten für die Forschung, Luftqualitätsdaten. Dies könne auch ein Dilemma sein, wie ein Beispiel auf dem Bereich IT-Sicherheit oder auch im Gesundheitswesen zeigt. Auf der einen Seite benötigen Datennutzende möglichst viele Nachweise zur Korrektheit der Daten, auf der anderen Seite fürchten Datengebende einen Reputationsverlust oder Verluste im Schutz persönlicher Gesundheitsinformationen.

Anreize für Datennutzende lägen im eigenen Mehrwert von bereitgestellten Daten anderer zur Weiterentwicklung oder Optimierung von Produkten/Dienstleistungen, im Beitrag zum Gemeinwohl, in der Haftungsreduktion durch den DT (Medizin), oder in Anreizen über direkt erfahrbare/sichtbare Mehrwerte (Beispiel PayBack).

7.5 Zusammenfassung und Ausblick

Die größten **Hemmnisse bei der Etablierung von Geschäftsmodellen** liegen den Erhebungsdaten zufolge in der fehlenden Bereitschaft, Daten zu teilen. Hier besteht eine Herausforderung seitens DT-Betreiber, finanzielle oder nicht finanzielle Anreize zu schaffen, um eine höhere Bereitschaft zu erwirken. Auf der anderen Seite deuten die Erhebungen auf eine hinreichende Nachfrage am Markt seitens der Datennutzenden.

Laut Online-Befragung bestehen zudem Herausforderungen hinsichtlich Finanzierungs- und Monetarisierungsansätze. Hier werden vor allem **fehlende Datenbewertungsstandards und Instrumente zur Datenbepreisung** als hinderlich gesehen, aber auch die Erwartungen fehlender oder geringer Wirtschaftlichkeit der DT sowie die Entwicklung eines Modells zu deren nachhaltiger Finanzierung. Diese Befunde bestätigen die Ergebnisse der Literaturanalyse in AP1.1. Während die Onlinebefragung suggeriert, dass fehlendes Know-How und fehlende Netzwerke keine übergreifenden Hindernisse (potenzieller) DT-Betreiber darstellen, scheint dies den Interviews zufolge in Einzelfällen durchaus eine Rolle zu spielen.

Als **Organisationsform** für DT wählte die große Mehrheit der Pilotprojekte öffentliche Einrichtungen (zum Beispiel Hochschulen). Dies ist im Hinblick auf die befragte Gruppe (öffentliche Fördernehmer) nicht unbedingt überraschend und als Befund entsprechend einzuordnen. Grundsätzlich sprachen sich im Kontext der Interviews und der Fachgruppenworkshops auffällig viele Personen (wenn auch nicht alle) für eine staatliche/öffentliche DT-Lösung im Kontext besonders sensibler Daten, insbesondere Gesundheitsdaten, aus. In der Diskussion während der Fachgruppenworkshops wurde zudem das Modell der Genossenschaft positiv hervorgehoben.

Hinsichtlich der langfristigen Finanzierung nach Förderende sehen die meisten Pilotprojekte vor allem eine Finanzierung durch Datennutzende vor, zum Beispiel durch Gebühren für den Zugriff auf Daten oder für Dienstleistungen. Ein deutlicher Anteil der Pilotprojekte plant, trotz vermuteter Innovationskraft gegenüber privaten Anbietern, mit einer öffentlichen Finanzierung.

Aufgrund des derzeitigen Entwicklungsstands der Pilotprojekte zeichnet sich noch kein deutlicher Trend hinsichtlich **Zahlungsmodalitäten und Bepreisung** ab. Entsprechend ist dieses Thema in der weiteren Begleitforschung vertiefter zu beleuchten.

Bislang liegt **keine deutliche Evidenz einer zukünftigen Monopolbildung am Markt** vor. Aufgrund der in diesem Kontext vorliegenden Netzwerkeffekte sei dies jedoch zu einem gewissen Grad zu erwarten. Ob eine Monopolbildung in diesem Fall grundsätzlich problematisch ist, wurde unter Interviewpartnerinnen und -partnern sowie unter den Pilotprojekten unterschiedlich gesehen. Neben dem Risiko der Monopolbildung könnten zukünftig auch kartellrechtliche Risiken eine Rolle spielen. Letzteres bleibt zunächst eine offene Forschungsfrage.

8 QT4: Akzeptanz, Skalierung und Transfer

QT4 behandelt mehrere Unterthemen: Akzeptanz, Skalierung, Standardisierung, Zertifizierung und Akkreditierung, sowie die Rolle staatlicher Infrastrukturen und Förderung bei der Etablierung von DTM. Für jedes Unterthema werden mehrere Forschungsfragen untersucht. Zur Beantwortung der Fragen wurden jeweils die Zwischenberichte der geförderten Projekte gesichtet und Interviews mit deren Repräsentanten sowie externen Experten geführt. Ferner wurde eine Online-Umfrage unter den Förderprojekten durchgeführt und im Vorfeld der Interviews eine Auswertung der bestehenden Forschungsliteratur vorgenommen (Bericht 1.2). Schließlich wurden die bis dahin vorliegenden Zwischenergebnisse am Fachgruppenworkshop am 7. September 2023 mit interessierten Teilnehmenden aus den Förderprojekten diskutiert. Die im Folgenden dargestellten Forschungsergebnisse basieren auf den so erhobenen Informationen.

8.1 Akzeptanz

QT4 versteht „Akzeptanz" als die Bereitschaft von (potenziellen) Datengebenden und Datennutzenden, am DT teilzunehmen. Hierzu wurden folgende Forschungsfragen gestellt:

Welche Faktoren beeinflussen die Akzeptanz von DT bei Datengebenden wie Datennutzenden?

Welche Rolle spielen insbesondere Datensouveränität, Mitwirkung und FRAND/FAIR-Bedingungen für die Akzeptanz?

Wie versuchen DT, Akzeptanz zu erreichen? Welche Strategien und Maßnahmen haben sich als erfolgreich erwiesen? Wo gibt es Probleme und Herausforderungen?

Grundsätzlich wird Akzeptanz (zusammen mit Skalierung und Transfer) von etwas weniger als der Hälfte der Förderprojekte als (eher oder sehr) hohes Hemmnis gewertet. Nur drei Projekte sehen hier „eher kein" Hemmnis. Damit sind die Hemmnisse hier etwas weniger verbreitet als bei Geschäftsmodellentwicklung (genau die Hälfte) und bei den rechtlichen Rahmenbedingungen (über Zweidrittel).

Die Förderprojekte verorten **Akzeptanzprobleme** vor allem bei den Datengebenden, weniger bei Datennehmenden. Unzureichende Bereitschaft, Daten zu teilen wird von 13 der 15 Projekte, als Hemmnis gesehen (sechs bewerten dieses Hemmnis als: „sehr starkes"); unzureichende Nachfrage nach Daten von Seite potenzieller Nutzenden hingegen nur von zwei Projekten. Diese Einschätzung ist plausibel, insofern als dass der Nutzen geteilter Daten tendenziell eher bei den Nutzenden, die Risiken aber eher bei den Datengebenden akkumulieren dürften (siehe weiter unten)

Die Interviews, Umfrage, Workshops und Fortschrittsberichte („Quellen") indizieren, dass vier wesentliche Faktoren die Akzeptanz beeinflussen: **Sicherheit/Vertrauen**, **Nutzen**, **Kosten/Aufwand** und **Altruismus**.

8.1.1 Sicherheit und Vertrauen

Daten zu teilen, birgt für Individuen wie Firmen Risiken. Personenbezogene Daten können illegitim ausgewertet werden, Geschäftsgeheimnisse an Wettbewerber abfließen, und Compliance-Verstöße des DT oder der Datennutzenden auf die Datengebenden zurückfallen. Diese Missbrauchsrisiken möglichst auszuschließen und bei Datengebenden Vertrauen aufzubauen, dass diese Risiken minimiert worden sind, wurde in allen Quellen übereinstimmend als wichtiger Treiber von Akzeptanz beschrieben. Etwas weniger stark betont, aber nicht unwichtig ist, dass die **Nutzung** geteilter Datennutzung ebenfalls Risiken bergen kann: inkorrekte

Daten oder Metadaten (z.B. Einheiten) können gefährliche Fehler auslösen, Compliance Versagen bei den Datengebenden oder dem DT zu Rechtsverstößen auch auf Seite des Nutzenden führen, und Aufwände sich schlicht nicht rentieren, wenn Daten sich als weniger nützlich denn erhofft erweisen.

Verschiedene Maßnahmen wurden in den Quellen benannt, um Sicherheit und Vertrauen zu gewährleisten bzw. aufzubauen. Die Gewährleistung von **Datenschutz und Datensicherheit** wurde übereinstimmend als von zentraler Wichtigkeit genannt. Ähnliches gilt für **Datensouveränität**, d.h. das Datengebende stets einsehen und kontrollieren können, wer ihre Daten für was nutzt, und im Zweifel Nutzungen untersagen können. **Transparenz** ist somit kritisch, da sonst keine Datensouveränität praktisch möglich ist, und wird in den Quellen ebenfalls entsprechend stark betont.

Hervorzuheben ist, dass **Datenschutz, Datensicherheit und -souveränität (DS3)** keineswegs nur Themen für Datengebende sind: Im Gegenteil schätzen die Förderprojekte die Rechtssicherheit der Datennutzung, welche DS3 (mit-)herstellt, als einen der wichtigsten Akzeptanzfaktoren bei Datennutzenden ein.

Gleichwohl steht DS3 tendenziell in einem Spannungsverhältnis zum ebenfalls sehr wichtigen Faktor **Aufwand/Kosten** (siehe unten): maximiert man DS3, steigen oft auch Aufwand/Kosten für Datengebenden wie -nutzende. Was das optimale Verhältnis von DS3 zu Aufwand/Kosten ist, kann zwischen verschiedenen Datengebenden variieren. Manche Datengebende können unter Umständen eine (etwas) niedrigere DS3 präferieren, wenn so auch Aufwand/Kosten minimiert werden. Während es datengebenden Firmen oft ausgesprochen wichtig ist, dass sie jede Nutzung bzw. Weitergabe ihrer Daten einzeln freigegeben müssen (und so volle Kontrolle und Transparenz über diese behalten), kann Gleiches für datengebende öffentliche Stellen eher eine Belastung darstellen. Für Letztere kann es wichtiger sein, dass die rechtliche Zulässigkeit und Compliance der Datenverarbeitungen seitens der Datennutzenden grundsätzlich garantiert ist, als eine (aufwändige) Datensouveränität über jede einzelne Verarbeitung auszuüben. Empirische Untersuchungen zu *medical consents* mit Patientendaten haben ebenfalls ergeben, dass Menschen Datenschutz zwar grundsätzlich wichtig ist, sie aber mitunter Lösungen präferieren, die Abstriche bei DS3 vorsehen, um den Aufwand für sich selber als Datengebende zu begrenzen.

Die praktische Umsetzung von DS3 erfolgt i.d.R. über technische, rechtliche und organisatorische Maßnahmen. Hier gibt es viele Optionen. Hervorzuheben ist, dass die Förderprojekte – wie auch die interviewten Expertinnen und Experten – **technische Architekturen** für am besten geeignet hielten, um Akzeptanz zu stiften. Hervorzuheben ist, dass bei diesen Architekturen die Daten entweder direkt zwischen Datengebenden und -nutzenden geteilt werden (*p2p*, ohne „Umweg" über den DT) oder erst gar nicht die IT-Systeme der Gebenden verlassen, sondern stattdessen vor Ort anhand der Algorithmen oder Fragen der Nutzenden verarbeitet werden (*algorithm to data*).,. Auch **Zertifizierungen** – sowohl des DT als auch der Datengebenden oder -nutzenden – wurden von Experten wie Förderprojekten als probates Mittel identifiziert, um DS3 zu sichern und zu demonstrieren.

Die **Reputation** der Datengebenden und -nutzenden sowie des DT kann aus Sicht der Förderprojekte einen begrenzten Beitrag zur Akzeptanz leisten. Zwölf Förderprojekte bewerten die **Reputation der Datennutzenden** als grundsätzlich wichtig für die Akzeptanz des DT bei den *Datengebenden*; drei Projekte halten diese jedoch für eher nicht wichtig. 15 Förderprojekte sehen die **Reputation des DT** als wichtig für die Akzeptanz der Datengebenden; keines hält diese für nicht wichtig. Gebeten, die *zwei wichtigsten* Faktoren für die Akzeptanz des DT bei den Datengebenden auszuwählen, nennt allerdings *kein* Förderprojekt die Reputation der

Nutzende oder des DT. Demzufolge scheint Reputation aus Sicht der Projekte nicht unwichtig, aber auch nicht ausschlaggebend zu sein.

Ein ähnliches Bild bietet sich bei der Frage zum Beitrag von Reputation für die Akzeptanz des DT bei den *Datennutzenden*. Sechs Projekte glauben, dass die **Reputation der Datengebenden** für Akzeptanz bei den Datennutzenden zumindest „eher wichtig" ist; drei Projekte geben jedoch an, dass diese „eher nicht" wichtig sei. Die **Reputation des DT** ist hingegen für elf Projekte wichtig für seine Akzeptanz bei den Datenutzenden, nur ein Projekt sieht sie als „eher nicht" wichtig. Nach den *zwei wichtigsten* Faktoren für die Akzeptanz des DT bei den Datennutzenden gefragt, nennt jedoch nur ein Projekt die Reputation des DT, und *keines* die Reputation der Datengebenden.

Die **Neutralität des DT** gilt ebenfalls als kritisch für die Akzeptanz. Konsens ist, dass die Interessen des DT beziehungsweise seines Betreibers nicht mit denen der Datengebenden und Nutzenden kollidieren dürfen. Diese Neutralität muss ferner für Datengebende und -nutzende klar und leicht ersichtlich dargelegt sein. Weniger Konsens besteht in der Frage, welche praktischen Implikationen sich daraus ableiten, etwa was die Identität des Betreibers, seine geschäftlichen Aktivitäten und die (non-) Profit-Orientierung des DT betrifft.

Die große Mehrheit der Förderprojekte glaubt, dass es für die Akzeptanz des von ihnen entwickelten DT am besten wäre, wenn sein Betrieb langfristig von einer staatlichen Stelle (elf Projekte), öffentlichen Hochschule oder Forschungseinrichtung (zwölf Projekte), altruistischen Gesellschaft (ein Projekt) oder Genossenschaft (ein Projekt) übernommen wird. Privatwirtschaftliche Akteure mit Profitabsicht dagegen werden nur von einem Projekt für geeignet gehalten, um die Akzeptanz zu fördern.[1] Ebenso skeptisch werden Profitabsichten gesehen: zwölf von 16 Projekten glauben, dass es für die Akzeptanz „sehr wichtig" ist, dass der DT „auf non-profit Basis betrieben wird". Nur vier der Förderprojekte sehen einen kommerziellen Betrieb ihres DT vor.

Diese Präferenz für staatliche Stellen als Betreiber wird von den interviewten Experten überwiegend nicht geteilt. Im Gegenteil sehen diese staatliche Stellen hier skeptisch, und eine Profitorientierung als DT nicht zwangsläufig als problematisch an, sofern das Geschäftsmodell den DT nicht zum Wettbewerber der Datengebenden oder -nutzenden macht. Eine mögliche Erklärung für diese divergierenden Einschätzungen könnte darin liegen, dass alle drei der interviewten Experten in engem beruflichem Kontakt zur Privatwirtschaft stehen, während viele Projektvertreter eher aus dem akademischen bzw. medizinischen Umfeld stammen.

Im Zusammenhang mit der Frage der **Neutralität** wurde auch die Frage der Auswirkungen auf Akzeptanz untersucht, wenn der DT zusätzliche Dienste anbietet (z.B. Datenauswertung, Beratung, Unterstützung bei sonstiger Daten-Compliance oder Digitalisierung). Der *Data Governance Act* scheint Dienstleistungen, die über reine Datenvermittlung (data intermediation) und assoziierte Datenaufbereitung (z.B. Pseudonymisierung) hinausgehen, strikte Grenzen zu setzen, insofern er die wirtschaftliche Nutzung von Daten durch den DT-Betreiber untersagt. Die Entwicklung und Ausführung von Zusatzdiensten würde aber eine

[1] Weitere mögliche Betreiber, die allerdings niedrigere Zustimmungswerte erfahren, sind Verbände (zwei Projekte) und Konsortien wichtiger Branchenakteure wie Firmen oder Krankenhäuser (ein Projekt). Trotz dieser schlechten Einschätzung von Verbänden und Firmenkonsortien als Betreiber *für die Akzeptanz* des DT, geben sieben Projekte an, dass sie Verbände als Betreiber für ihren DT vorsehen bzw. „aktuell [für eine der] sinnvollsten Optionen" halten (Firmenkonsortien: sechs Projekte; privatwirtschaftliche Akteure: 4 Projekte). Eine mögliche Erklärung ist, dass die Projekte bei der Beurteilung der „Betreiber-Fähigkeit" von Akteuren die Maximierung der Akzeptanz des DT zwar als ein wichtiges, aber nicht ausschließliches Kriterium sehen.

eigene wirtschaftliche Nutzung der Daten durch den DT implizieren.[2] Der Gesetzgeber scheint zu befürchten, dass wirtschaftliche Nutzung der Daten durch den DT grundsätzlich seine Neutralität in Frage stellen und damit Akzeptanzprobleme schaffen würde.

Die Förderprojekte und die interviewten Experten sahen diesen Punkt differenzierter. Während sie, wie oben besprochen, Neutralität als kritisch erachteten, bewerteten sie das Angebot zusätzlicher Dienstleistungen durch den DT überwiegend positiv. Zehn Förderprojekte glauben, dass zusätzliche Dienstleistungen die Akzeptanz des DT bei Nutzenden steigern würden. Nur eines befürchtet, dies könnte die Akzeptanz bei Datengebenden schwächen.[3]

8.1.2 Nutzen, Kosten und Aufwand

(Vertrauen in) die Sicherheit des DT ist Vorbedingung für seine Akzeptanz, schafft aber noch keinen positiven Anreiz, den DT zu nutzen. Die Interviews mit Experten wie Förderprojekten sowie die Gespräche beim Workshop unterstrichen, dass Akzeptanz im Sinne aktiven Datengebens bzw. aktiver Datennutzung, mit der **Wertigkeit der Use Cases** für die Beteiligten steht und fällt, denn daraus entsteht letztlich der Wert des DT für die Beteiligten. Dieser Wert muss zumindest perspektivisch größer sein als Aufwand und Kosten. Vor diesem Hintergrund wird verständlich, warum das Angebot von Zusatzdiensten durch den DT, die über reine Datenvermittlung und -aufbereitung hinausgehen, akzeptanzfördernd sein kann (s.o.): Zusatzdienste können zusätzliche oder höherwertige Datennutzungen ermöglichen, ohne dass die Nutzenden erst eigene Kompetenzen teuer ausbilden müssen, oder senken andere mit Datenteilen und –nutzen assoziierte Kosten (z.B. Compliance). Je niedriger die Kosten und Aufwände, desto eher kann der DT einen positiven Wert ermöglichen.

Kostenpunkte wurden nicht im Detail untersucht, dürften aber vor allem Personalkosten plus etwaige Gebühren für Datenzugang einschließlich Kompensation des Datengebenden sein. Personalaufwand fällt vor allem durch die Arbeitszeit an, die qualifiziertes Personal aufbringen muss, um Daten bereitzustellen und zu teilen, Use Cases und Datenbestände zu identifizieren und auf ihre Qualität zu prüfen, Übereinkünfte mit Datengebenden bzw. -nutzenden zu schließen, Daten aufzubereiten, Compliance zu gewährleisten, und schließlich die Datenverarbeitungen vorzunehmen. Diese und weitere Aufwände gering zu halten, wird von allen Quellen als kritisch für die Akzeptanz gesehen und tendenziell wichtiger als die Höhe etwaiger direkter Kosten wie Zugangsgebühren.[4]

Insbesondere die interviewten Expertinnen und Experten beschrieben die Identifikation wertvoller Use Cases als sowohl essentiell wie nicht-trivial. Ihnen zufolge erfordert die Identifikation und belastbare Entwicklung und Umsetzung wertvoller Use Cases oft intensive, zeitaufwendige Kommunikation und gemeinsame Arbeit zwischen den beteiligten Datengebenden und -nutzenden. Das schafft ein „Henne-Ei Problem“: der Aufwand lohnt sich nur, wenn der resultierende Use Case entsprechend wertvoll ist. Bevor der Aufwand aber investiert wurde, kann es schwierig sein, seine Wertigkeit abzuschätzen. Identifikation und

---

[2] Die juristische Auslegung scheint hier noch im Fluss zu sein.

[3] Im Workshop erwähnte ein Förderprojekt, dass avisierte Datengebende sogar gefragten hatten, ob der DT ihnen nicht allgemeine Digitalisierungsdienste anbieten könnte. In einem Interview sagte ein anderes Förderprojekt, dass gerade KMU starkes Interesse an solchen Zusatzdiensten, z.B. bei der Datenauswertung, haben dürften.

[4] So werteten 13 Förderprojekte „niedrige technische und Verwaltungsaufwände“ als „eher“ oder „sehr“ wichtig für die Akzeptanz des DT bei Datennutzenden (16 beziehungsweise 14 bei Datengebenden). Aber nur neun hielten die Höhe der direkten Kosten wie Zugangsgebühren für eher (sieben Projekte) oder sehr (zwei Projekte) wichtig.

Entwicklung von Use Cases wird zusätzlich erschwert, wenn die beteiligten Datengebenden und -nutzenden keine bestehende Beziehung zueinander haben.

Mehreren der Expertinnen und Experten zufolge, müssen DT daher oft auch als **Broker und Matchmaker** agieren, d.h. potentielle Use Cases einschließlich der möglichen Datengebenden und -nutzenden identifizieren und die Akteure in Workshops oder ähnlichen Formaten ins Gespräch bringen, um Möglichkeiten systematisch zu erkunden. Das erfordert vom DT eine relativ tiefe Kenntnis der am DT beteiligten Akteure (Datenbestände, Kompetenzen, Geschäftsmodelle etc.) sowie relativ personal- und zeitaufwendige Arbeit, und ein gewisses Standing unter den Akteuren.

In Gesprächen und der Umfrage mit den Förderprojekten fanden diese Thesen eher bedingte Unterstützung. Im Workshop wurde ihnen mit Einschränkungen (siehe unten) durchaus zugestimmt. In der Umfrage gab allerdings nur ein Projekt an, dass „fehlende Use Cases in der Praxis (best practice)" ein Hemmnis für die „Etablierung eines tragfähigen [DT-] Geschäftsmodells" seien. Im Gegenteil war für zehn Projekte dies explizit kein Hemmnis. Da die interviewten Expertinnen und Experten längere Erfahrung mit dem Aufbau und Betrieb von DT haben (und in einem Fall selbst einen DT in einem sehr weit fortgeschrittenen Stadium leiten), sollte ihre Einschätzung wahrscheinlich sehr ernst genommen werden.

Wie kann dann die eher abweichende Einschätzung in der Umfrage erklärt werden? Zum einen ist zu konstatieren, dass sich die DT der Förderprojekte überwiegend noch in der frühen Aufbauphase befinden.[5] Die Herausforderung, ein ausreichendes Volumen wertiger Use Cases zu entwickeln und umzusetzen könnte daher noch nicht voll zu Tage getreten sein. Es ist auch möglich, dass das Umfrage-Item mehr als Frage nach best practice bei der Etablierung von DT-Geschäftsmodellen verstanden wurde, denn als Frage nach der Herausforderung, wertige Use Cases für die Datengebenden und -nutzenden zu entwickeln, verstanden wurde. Schließlich kristallisierten sich im Workshop auch Einschränkungen heraus, welche die obigen Zahlen ebenfalls erklären könnten.

Insbesondere für den unter den Förderprojekten sehr stark vertretenen Medizinbereich stellte sich im Workshop heraus, die obigen Thesen der Experten gelten nur bedingt: im Medizinbereich ist die Grundstruktur der Use Cases (Verwendung von Patientendaten zur medizinischen Forschung) klar definiert, so dass der Aufwand, lohnenswerte Use Cases (hier vor allem Forschungsfragen und Studiendesigns) zu identifizieren, vergleichsweise gering ist. Ähnliches gilt für DT, wo es vor allem darum geht, wohlbekannte und standardmäßig definierte Geschäftsprozesse zu digitalisieren.

In Summe dürfte der Aufwand, wertvolle Use Cases zu identifizieren am höchsten – und die Übernahme einer Matchmaker-Funktion durch den DT für die Akzeptanz am wichtigsten – bei jenen DT seien, die die Entwicklung gänzlich neuer und innovativer Datennutzungen anstreben und Datengebende sowie -nutzende aus unterschiedlichen, bisher eher getrennt operierenden Sektoren oder Domänen zusammenbringen wollen.

Kritisch für Datennutzende ist in jedem Fall die **Qualität der Daten und Metadaten**. 14 bzw. 15 der Förderprojekte bezeichneten diese als „eher" oder „sehr wichtig" für die Akzeptanz. Die

[5] Kein Förderprojekt hat den Punkt der Markteinführung erreicht und nur drei Projekte befinden sich in der Erprobung ihres DT. Umgekehrt sind 14 noch damit beschäftigt, dass Konzept ihres DT auszuarbeiten oder haben es erst kürzlich finalisiert.

Vielfalt der Datensätze wurde hingegen nur von neun als eher oder sehr wichtig bewertet.[6] Gründe hierfür könnten die praktische Herausforderung für Datennutzende sein, passende Datensätze zu identifizieren, und für Datengebende, wertvolle Datensätze überhaupt bereitzustellen. Ein „weniger, aber dafür bessere Daten"-Ansatz seitens des DT kann somit sinnvoll sein, zumindest in der Aufbauphase. Laut der Förderprojekte hemmten zum Teil Datenschutzvorgaben (Privacy by Design) die Erhebung wertvoller Datensätze; zum Teil ständen auch die mangelnde Digitalisierung potenzieller Datengebenden der Aufbereitung und Bereitstellung von Datensätzen im Weg. Grundsätzlich helfen Metadaten, die den Datensatz beschreiben, Datennutzenden, passende Datensätze für ihren angedachten Use Case zu finden. Gute Metadaten zu erstellen ist jedoch oft zeitaufwendig und erfordert Fachwissen. Zudem kann ein Datensatz oft anhand vieler verschiedener Eigenschaften und Dimensionen sinnvoll beschrieben werden. Es gibt selten „die" eine Beschreibung, die alle Eigenschaften des Datensatzes wiedergibt, die für mögliche Nutzende relevant sein könnten. Welche Eigenschaften relevant sind, hängt vielmehr vom Use Case ab, und sie sind selbst potenziell sehr heterogen. Schlüsselwort- oder kurztext-basierte Annotationen und Suchfunktionen sind daher nur bedingt hilfreich, gerade um sektorübergreifende Use Cases zu ermöglichen. Wie diese Problematik zu lösen ist, ist noch offen. Technische Hilfen (z.B. KI) sind oft noch nicht ausgereift genug; „händische" Unterstützung z.B. durch einen als Matchmaker agierenden DT trifft schnell auf Skalierungsgrenzen.

8.1.3 Altruismus und Instrumentelle Anreize zum Datenteilen

Während der **Nutzen von DT für Datennutzende** (zumindest theoretisch) offensichtlich ist, und nur von wenigen Förderprojekte als Problem für die Akzeptanz und den Aufbau von DT gesehen wird, ist es schwieriger, einen klaren **Nutzen für Datengebende** zu identifizieren. Mangelnde Bereitschaft, Daten zu teilen wird entsprechend von 13 Projekten als Hemmnis gesehen, nur zwei sehen dies als „eher kein" Hemmnis. Für viele Förderprojekte scheint es auch schwierig zu sein, positive Anreize für das Datenteilen zu schaffen. Zwar planen mindestens zehn Förderprojekte eine monetäre Vergütung für die Bereitstellung von Daten oder erachten diese als sinnvoll, und fünf bzw. sechs wollen Datengebenden über andere instrumentelle Vorteile wie Zugang zu Daten anderer Datengebenden oder zu freien oder kostenreduzierten Diensten kompensieren. Jedoch erachten Zwölf von 15 Projekten die Schaffung solcher finanziellen Anreize und 13 die Schaffung nicht-finanzieller Anreize als eine Herausforderung für den DT-Betreiber in der Praxis. Nur fünf setzen auf Altruismus (Datenspende) als Motivation für das Datenteilen. Zweidrittel der Förderprojekte bewerten allerdings Altruismus („gesellschaftlicher Nutzen") als wichtigen Anreiz für das Datenteilen; nur etwa die Hälfte glaubt, dass monetäre oder andere instrumentelle Vorteile wichtig sind.

Diese Ergebnisse werfen die Frage auf, wie Datenteilen stärker gefördert werden kann, und insbesondere auch, inwiefern die Förderprojekte die anscheinend wichtige altruistische Motivation stärker bespielen können.

8.1.4 FRAND und FAIR Bedingungen

Die FRAND (*fair, reasonable, and non-discriminatory access*) und FAIR (*findable, accessible, interoperable, reusable*) Prinzipien spielen für die Förderprojekte zum Teil eine Rolle. Sieben

[6] Die Kluft zwischen den Gewichtungen von *Daten-/Metdatenqualität* und *Vielfalt der Datensätze* ist sogar noch größer als diese zusammenfassenden Zahlen anmuten lassen. Von den neun Förderprojekte, die „Vielfalt" als eher/sehr wichtig bewerteten, werteten 8 die Datensatzvielfalt nur als „eher wichtig", und 1 Projekt bewertete sie als „sehr wichtig". Umgekehrt werteten 11 Förderprojekte Datenqualität als „sehr wichtig", nur drei als „eher wichtig". (Metadaten: 7 Projekte „sehr wichtig", 8 „eher wichtig).

Projekte gaben an, dass der Datenzugang bei ihrem DT gemäß dem FRAND Prinzip erfolgt oder erfolgen soll, neun weitere machten entweder keine Angaben hierzu (sechs Projekte) oder indizierten, dass FRAND für sie keine Rolle spielte, da Datenzugangsentscheidungen letztlich nicht von ihnen beziehungsweise dem DT, sondern von den Datengebenden getroffen würden (drei Projekte). In diesem Kontext wurde auch auf die Spannung zwischen FRAND (*non-discriminatory access*) und dem Prinzip der Datensouveränität verwiesen. Ein Punkt, der auch in einem der Experteninterviews unterstrichen wurde.

FAIR wird von vier der Förderprojekte laut Umfrage implementiert, zehn weitere orientieren sich daran. In Interviews wurde wiederholt darauf ausgeführt, dass FAIR eigentlich für den Open Data Bereich gemünzt wurde, weshalb es zwar nützliche Orientierung biete, aber nicht immer direkt übertragbar sei.

8.2 Skalierung

Hinsichtlich der **Skalierung ihrer DT**, scheint die Mehrheit der Förderprojekte zumindest vorerst auf Deutschland und die aktuell vom DT adressierte(n) Branche(n) fokussiert zu sein. Nach dem mittelfristigen Ziel für ihren DT gefragt, gaben zehn Projekte an, dass sie hofften, dass „alle oder die meisten" Akteure in den von den DT adressierten Branchen am DT als Datengebende oder -nutzende teilnehmen sollten. Nur ein Projekt wollte auch mittelfristig „eher klein bleiben" und sah keinen Nachteil darin, sollten zahlreiche Branchenakteure *nicht* an ihrem DT teilnehmen. Umgekehrt hofften ebenfalls nur zwei Projekte, mittelfristig alle oder die meisten Akteure der adressierten Branchen in ganz Europa für eine Teilnahme am DT zu gewinnen.

Hinsichtlich der **Geschwindigkeit**, mit **der Skalierung** verfolgt werden sollte, vertreten zehn Förderprojekte die Position, dass eine eher langsame Skalierung, „graduell und über einen längeren Zeitraum", unproblematisch sei. Vier Projekte hingegen sind der Auffassung, die Skalierung sollte „möglichst rasch" angegangen werden.

Ein möglicher Grund, warum die Förderprojekte – im Gegensatz zu klassischen Digital-Startups aus dem B2C Sektor – Skalierung überwiegend für eher graduell angehbar halten, könnte in ihrer Wahrnehmung der **Marktstruktur und Wettbewerbssituation** liegen. Laut der Umfrage erwartet die große Mehrheit der Projekte, dass sich eine oligopolistische oder sogar eine monopolistische Wettbewerbsstruktur, mit nur einer kleinen Zahl bzw. einem einzigen DT pro Branche oder Anwendungsbereich herausbilden wird. Gleichzeitig sehen sich viele der Förderprojekte mit keinen Wettbewerbern konfrontiert. Die, die Wettbewerber identifiziert haben, nennen meist staatliche oder halbstaatliche Projekte (z.B. Elektronische Patientenakte, Forschungsdatenzentrum, European Health Data Space). Diese werden möglicherweise als weniger aggressiv wahrgenommen als Konkurrenten aus der Privatwirtschaft. Wenn man aber vor einer oligopolistischen Wettbewerbsstruktur ausgeht und keine oder sehr wenige Konkurrenten am Markt sieht, kann es mindestens rational sein, Skalierung eher langsam anzugehen. Hierfür spricht auch, dass die „Kunden" der DT der Förderprojekte (Datengebende und -nutzende) meist Unternehmen, öffentliche Stellen oder Krankenhäuser sind, die DT also primär im B2B-Geschäft stehen. Im Gegensatz zum B2C ist extrem rasche Skalierung im B2B-Sektor oft weniger wichtig, von Anfang an mit belastbaren Angeboten in den Markt zu treten dafür aber wichtiger.

Als kritischer denn die Frage der Geschwindigkeit könnte sich der Fokus auf Deutschland vieler Projekte erweisen. So konstatierte einer der interviewten Experten, dass die Teilnehmende an DT, zumindest in der Privatwirtschaft, in der Regel global, „mindestens aber europäisch", dächten. Inwiefern ein primär Deutschland-fokussierter DT für solche Akteure langfristig interessant sein kann, ist offen.

Gleiches gilt für weitere Fragen hinsichtlich Skalierung, angefangen damit, wie genau diese zu verstehen ist. Die Förderprojekte, so der Eindruck aus den Interviews und dem Workshop, scheinen Skalierung als die Zahl der teilnehmenden Datengebenden und -nutzenden zu verstehen. Relevante Metriken könnten aber auch die Zahl der Sektoren oder sogar der anderen DT sein, an die der eigene angebunden ist, die Vielfalt der vom DT verwalteten Datensätze, oder das Transaktionsvolumen (Menge der geteilten Daten, Zahl der Datenaustausche). Wie Skalierung praktisch angegangen werden kann (z.B. wie Zusammenschlüsse oder Interoperabilität über Branchengrenzen hinweg gesichert werden kann) und welche Herausforderungen und trade-offs sich bei der Skalierung stellen (z.B. Fähigkeit, zeitintensives persönliches Matchmaking zwischen Akteuren zu betreiben) sind weitere Fragen.

8.3 Standardisierung, Zertifizierungen und Akkreditierungen

Ausgangslage der Untersuchungen in AP1.2. in QT4 waren die folgenden Forschungsfragen:

Was ist der Stand bei Standards und Zertifizierungen? Welche Standardisierungsaktivitäten wurden schon initiiert beziehungsweise welche Standards wurden schon entwickelt? Gibt es hinreiche Anreize für die Entwicklung, aber auch die Implementierung von Standards?

Wie wichtig für den Aufbau von DT sind Standards beziehungsweise Standardisierung zu diesem Zeitpunkt in den folgenden Feldern: Begriffe? IT-Technik? Daten? Governance einschließlich Musterverträge (standard contractual clauses)?

Welche Vor- beziehungsweise Nachteile haben Standardisierung als selbstregulierende Maßnahme?

Welche Rolle spielt Zertifizierung von DT für den Abbau von Informationsasymmetrien? Welche Anreize haben Zertifizierer, um Zertifizierungen für DT zu entwickeln und anzubieten?

Welche Qualitäten muss ein Zertifizierer nachweisen (z.B. im Rahmen seiner Akkreditierung), um aus Sicht der Akteure vertrauenswürdig zu sein?

8.3.1 Rolle von Standards und Standardisierungsaktivitäten

Aktuell gibt es bis auf den ISO-Standard ISO 20387 zu „General requirements for biobanking“, welcher Anforderungen an die Kompetenz, die Unparteilichkeit und den kohärenten Betrieb von Biobanken definiert, keinen international anerkannten Standard spezifisch zur DT. Ferner ist dieser Standard selbst unter etablierten Biobanken nicht bekannt. Parallel werden aktuell nur sehr wenige Standardisierungsaktivitäten speziell zu DT vorangetrieben. Laut Zwischenberichte aber auch auf Basis der Aussagen in den Interviews spielen für die große Mehrheit der geförderten Projekte der Einsatz und somit die noch zu leistende Entwicklung von Standards meist noch eine untergeordnete Rolle. Jedoch wird von der großen Mehrheit der geförderten Projekte die Interoperabilität der Daten als eine wichtige Voraussetzung für die Skalierung ihres DT und für die generelle Schaffung von Akzeptanz wahrgenommen. Hierbei wird jedoch die wichtige Rolle von Standards höchstens indirekt rezipiert.

Insbesondere werden Datenstandards, inkl. Formate, Ontologien, Einheiten, Metadaten, genannt, um den Austausch und die Nutzung von Daten vereinfachen und forcieren zu können. Ferner werden Standards für das Qualitätsmanagement der Datenverarbeitung von Seiten der Datennutzenden begrüßt. So sollte es bestimmte Grundstandards geben, z.B. um den Datenzugriff zu protokollieren oder die Auditierbarkeit des Datenflusses zu ermöglichen, um eine Nachvollziehbarkeit der Qualität der Datenverarbeitung nach außen hin auch zu gewährleisten.

Hierzu werden verschiedene, meist bottom-up-getriebene **Standardisierungsaktivitäten** vorangetrieben, die aber meist nicht sektor- oder domänenübergreifend oder gar international koordiniert werden. Diese Fragmentierung wird damit begründet, dass Datenstandards häufig branchenspezifisch oder sogar projektspezifisch ausgearbeitet werden müssen, was konzertierte Standardisierungsinitiative innerhalb und über Branchengrenzen hinweg nicht sinnvoll erscheinen lassen. Hier ist auch anzumerken, dass nur wenige DT den Anspruch haben, demnächst auf globaler Ebene aktiv zu werden, wofür dann auch internationale Standards unabdingbar wären. Jedoch haben einige Projekte bereits branchenspezifische Standards umgesetzt bzw. sich grob daran orientiert. Im Medizinbereich handelt es sich um FHIR (Fast Healthcare Interoperability Resources) zum elektronischen Austausch von Gesundheitsdaten, in der Forstwirtschaft um StanForD, einen Standard für die Kommunikation zwischen Computern in Forstmaschinen, und in der Pflanzenzüchtung um die Breeding API (BrAPI), eine Schnittstelle für den Austausch von Pflanzenphänotyp- und -genotypdaten zwischen Pflanzenzuchtanwendungen, und MIAPPE, einen offenen Datenstandard zur Harmonisierung von Daten aus Pflanzenphänotypisierungsexperimenten. Weiterhin sieht die Mehrheit der geförderten Projekte einen Bedarf an Standards für die DT-Governance (Organisation, Datenzugangsbedingungen, Musterverträge) und -Infrastruktur (Schnittstellen, Sicherheitslevel). Jedoch hat sich dieser Bedarf aktuell noch wenig konkretisiert.

Komplementär zu diesen branchen- oder domänenspezifischen Standards werden **branchenübergreifende internationale Standards**, wie ISO 27001, für die Datensicherheit oder zur Softwarenentwicklung vor allem von den Projekten genutzt, die sich schon näher an der Markteinführung befinden. In diesem Kontext muss auch noch das Resource Description Framework (RDF), ein Standard des World Wide Web Consortium (W3C) genannt werden. Er spezifiziert ein Datenmodell für Metadaten und stellt auch eine wichtige Schnittstelle zu GAIA-X dar. Dasselbe gilt für den W3C-Standard (Open Digital Rights Language), der auch von GAIA-X empfohlen und vom International Data Space (IDS) schon genutzt wird. In diesem Kontext ist auch DIN SPEC 27070 („Anforderungen und Referenzarchitektur eines Security Gateways zum Austausch von Industriedaten und Diensten“), die im Rahmen von KomDatIS, der Datenplattform des Daten-Kompetenzzentrums für Städte und Regionen (DKSR) GmbH, eine einheitliche und standardisierte Datenstruktur ermöglicht. Hierbei wird auch die Weiterentwicklung und Integration von „Trust Frameworks“ sowohl der International Data Spaces Association (IDSA) als auch Gaia-X in das Open Source Framework Eclipse Dataspace Components (EDC) erwähnt. Dies unterstreicht die wichtige, aber bisher grundsätzlich noch wenig ausgenutzte Schnittstelle zwischen Standards und Open Source, die möglicherweise auch für die Weiterentwicklung von DT eine Rolle spielen könnte.

Neben der grundsätzlich positiven Wahrnehmung von Standards wurde kritisch angemerkt, dass bei einer umfassenden Interoperabilität aller Daten eines DT das Risiko des Datenmissbrauchs steigt, da nun ohne größeren Aufwand Daten verknüpft und damit unter anderem de-anonymisiert werden können. Dies stellt eine Analogie zu Gegenmaßnahmen im Maschinenbau dar, wo bewusst Abweichungen von etablierten Standards, eine so genannte Destandardisierung, in Kauf genommen werden, um ein Reengineering von hochwertigen Technologien, vor allem Maschinen zu verhindern. Diese Problematik stellt bisher eine Einzelmeinung dar.

8.3.2 Rolle von Zertifizierungen und Akkreditierung

Hinsichtlich der **Zertifizierung**, welche als wichtig für die Schaffung von Akzeptanz und Vertrauen von DT aber weniger relevant zum Abbau von Informationsasymmetrien zwischen Datengebenden, Datennutzenden und anderen Akteuren angesehen wird, stellt sich zunächst die Frage, welche Organisation im Datenökosystem zertifiziert werden sollte. Zunächst wird hier

der DT genannt. Eine Zertifizierung von DT oder eine Kennzeichnung analog etwa zum CE-Kennzeichen wird von den geförderten Projekten als langfristig wünschenswerte vertrauensbildende Eigenschaft betrachtet, unter anderem auch um die Neutralität nachzuweisen und der breiten Öffentlichkeit Neutralität zu signalisieren. Derartige Zertifikate sind aber höchstens im Aufbau und werden von den Projekten selbst i.d.R. noch nicht direkt verfolgt. Als Ausnahme ist ePrivacy zu nennen, das sich im Rahmen von TreuMed intensiv mit der Entwicklung eines Geschäftsmodells für eine Datensicherheitszertifizierung speziell für Apps für die biomedizinische Forschung im FeatureCloud AI Stores beschäftigt hat. Darauf aufbauend wurden erste Schritte hin zur Konzeptionierung eines Datensicherheits-Siegel ePrivacyFC, das an datensicherheitskonforme Apps des Feature Cloud AI Stores vergeben werden soll, unternommen.

Neben dem DT selbst stellt sich die Frage, ob die Datengebenden und die -nutzenden auch zertifiziert werden sollten, um zum einen die Datenqualität zu sichern und zum anderen Missbrauch bei der Datennutzung zu verhindern. Diese Forderung wurde vereinzelt gestellt, jedoch wird dies noch nicht durchgeführt und auch noch nicht geplant. Es wurde auch kritisch angemerkt, dass eine Zertifizierung aller Teilnehmenden ein massives Hemmnis für die schnelle Skalierung der DT bedeuten würde. Wenn Zertifizierungen durchgeführt werden, dann auf Basis branchenspezifischer Datenstandards, wie RDF, oder allgemeiner IT-Sicherheitsstandards, wie ISO 27001.

Im Gegensatz zur grundsätzlichen positiven Bewertung von Datenstandards, aber auch spezifischen Standards für DT zusätzlich zu den etablierten IT-Managementstandards, wird die weitergehende Zertifizierung etwas ambivalenter wahrgenommen. Denn externe Zertifizierer können nach Einschätzung einiger Experten nicht das Vertrauen in DT schaffen, welches intern durch die vertrauensvolle Kooperation geschaffen werden. Einziger Nutzen externer Zertifizierung wäre die Interoperabilität mit anderen DT aus anderen Branchen zu vereinfachen.

Die **Akkreditierung** von Organisationen, wie dem TÜV, die Zertifizierungsprogrammen für DT und deren Datenmanagement anbieten, wird aktuell nur von einer kleinen Minderheit der geförderten Projekte als auch der interviewten Experten thematisiert. Hier stellt sich die Frage, ob mit der Akkreditierung eine weitere Institution geschaffen werden sollte, um das Vertrauen und die Akzeptanz von DT weiter zu stärken. Alternativ könnten dazu auch staatliche Institutionen, wie das Bundesamt für Informationssicherheit beitragen. Grundsätzlich spielt die Akkreditierung von Institutionen, die DT auf Basis anerkannter Standards zertifizieren, aktuell noch keine relevante Rolle für die geförderten Projekte.

8.4 Staatliche Infrastrukturen und Förderung

Folgende Forschungsfragen wurden zur Rolle des Staates gestellt:

Wie könnte der Staat beziehungsweise die Forschungs- und Innovationsförderung zur Akzeptanz von DT beitragen?

Welche Politikinstrumente sollten zur Akzeptanzsteigerung noch herangezogen werden?

Welches Potential bieten Ansätze wie GAIA-X für die Steigerung der Akzeptanz bzw. Skalierung von DT?

Wie oben diskutiert, sind die Positionen in der Frage ob staatliche Stellen als DT-Betreiber fungieren sollen, divers. Größere Einigkeit bestand darüber, dass DT – ob von staatlichen Stellen oder anderen Akteuren betrieben – in jedem Fall vor staatlichen Zugriffen auf die Daten geschützt werden müssen. Insofern ein DT Datenbestände zusammenführt und zugänglich macht, besteht immer auch das Risiko, dass staatliche Stellen (z.B. Sicherheitsbehörden, Steuerämter) versuchen, diese Daten für ihre eigenen Zwecke abzugreifen. Dieses Risiko wurde

in den Interviews mehrfach angesprochen und darauf hingewiesen, dass derartige Vorgänge Akzeptanz in hohem Maße gefährden könnten.

Über diese Thematik hinaus sahen die interviewten Projekte und Experten primär drei **Rollen für den Staat**: Förderung, Öffentlichkeitsarbeit und das Setzen von Rahmenbedingungen. Allgemeiner Konsens war, dass dem Staat auf absehbare Zeit eine wichtige Rolle als (finanzieller) Förderer von DT zukommen müsse, da die Entwicklung tragfähiger privatwirtschaftlicher Finanz- und Geschäftsmodelle ein schwieriger Prozess ist. So betrachten 14 der befragten Projekte staatliche Anschubfinanzierung als wichtig für die Entwicklung von DT; nur ein Projekt hält Anschubfinanzierung für eher unwichtig. Zehn Projekte wollen ihren DT auch langfristig (d.h. nach Ende der gegenwärtigen Projektförderung) über öffentliche Zuwendungen finanzieren.

Die mögliche Rolle und Modalitäten nicht-finanzieller Förderung durch den Staat wurden soweit nur gestreift, könnten aber in der nächsten Phase der Begleitforschung vertieft betrachtet werden. In Interviews mit Förderprojekten wurde wiederholt Öffentlichkeitsarbeit angesprochen, dass der Staat also die Idee von DT bewerben könnte, um ihre Akzeptanz zu steigern. Es könnten aber auch weitere Formen der Unterstützung denkbar sein.

Schließlich wurde dem Staat eine Rolle bei der Förderung von Standardisierung und Zertifizierung zugesprochen. Hier wurde die Rolle staatlich geförderter Infrastrukturprojekte wie Gaia-X, die International Data Spaces Association und die European Data Spaces betont, um Standards und Zertifizierungen voranzutreiben und in die Breite zu tragen.

8.5 Zusammenfassung und Ausblick

Der Fokus der bisherigen Untersuchungen lag vor allem auf Akzeptanz, Skalierung und Standardisierung und Zertifizierung. Die Frage staatlicher Förderung und Infrastruktur wurde ebenfalls angeschnitten. Verschiedene weiterführende Fragen haben sich aus der bisherigen Arbeit ergeben, die im kommenden Jahr weiterverfolgt werden können. Zu diesen zählen insbesondere die Frage, wie Datengebende angereizt werden können, an DT zu partizipieren und Daten zur Verfügung zu stellen. Hier stellt sich etwa die Frage, wie Datenaltruismus stärker gefördert und praktisch unterstützt werden könnte. Bei der Skalierung stellt sich unter anderem die Frage, wie branchenübergreifendes Wachstum und der Zusammenschluss mehrerer DT ermöglicht werden kann. Das Problem der internationalen Expansion in europäisches und außereuropäisches Ausland stellt sich ebenfalls, wie auch die Frage der Integration mit Initiativen wie Gaia-X und der European Data Spaces. Auch die Themen Standardisierung und Zertifizierung sind weiterzuverfolgen. Hier stellt sich insbesondere die Frage, ob es zu einer Konsolidierung der diversen einschlägigen Aktivitäten kommt, oder eher zu weiterer Fragmentierung. Schließlich könnten auch die Optionen für weitere staatliche Förderung, gerade auch nicht-finanzieller Art, weitere Betrachtung verdienen.

9 Schlussfolgerungen und Ausblick

Die in den Pilotprojekten erprobten DTM sind unterschiedlich weit vorangeschritten. In der Konzeptionierung zeichnen sich projektübergreifende zentrale Herausforderungen ab:

Bei der **technischen Architektur** zeichnen sich zwei unterschiedliche Ansätze ab, die unter den Begriffen „Data Space“ und „Datenintermediär“ je nach Zentralisierungsgrad zusammengefasst werden können. Für Interoperabilität von Datensätzen erforderliche Schnittstellen sind noch in Entwicklung, wobei Datenschutzbedenken und Interessenkonflikte zwischen Datengebenden und -nutzenden diese ausbremsen. Technischen Maßnahmen wie Pseudonymisierung, asymmetrischer Kryptographie und mehrstufigen Zugriffs- und Nutzungskontrollen spielen eine wichtige Rolle bei der Gewährleistung von Datensicherheit und -souveränität. Im weiteren Verlauf der Begleitstudie wird ein Fokus auf das Zusammenwirken zwischen den technischen Bausteinen und den rechtlichen Rahmenbedingungen und der Entwicklung gemeinsamer Schnittstellen und standardisierter Datenformate liegen.

Aus **rechtlicher Perspektive** hat die Datenerhebung ergeben, dass sich die Annahme eines Wert-Risiko-Dilemmas mit der Aufgabe für DT, Compliance-Risiken zu senken und somit Datenteilen zu inzentiveren, als theoretischer Ansatz bewährt hat. Im weiteren Studienverlauf sollen drei Modelle des Datenteilens und der Aufgaben für Datenintermediäre weiter auf ihre rechtlichen Auswirkungen hin untersucht werden: Ein Modell „offener Daten“, bei der sich die Aufgaben für Intermediäre auf Organisatorisches beschränken, ein Modell ‚geteilte Daten‘ bei dem ein Intermediär (wie beispielsweise ein DT) Nutzungsberechtigungen prüft und Vereinbarungen trifft und diese technisch unterlegt, und ein Modell „Geteilte Analyseergebnisse“ bei welchem der DT als Datenverarbeiter auftritt und dementsprechend höheren rechtlichen Anforderungen gerecht werden muss. Diese Modelle werden im weiteren Begleitforschungsprozess konsolidiert und ausdifferenziert.

Aus **Geschäftsmodellsicht** liegen Hemmnisse für die erfolgreiche Etablierung von DT in der Anreizung von Datenteilen, wohingegen eine ausreichende Nachfrage seitens potenzieller Datennutzenden bestünde. Fehlende Datenbewertungsstandards und -instrumente zur Datenbepreisung erschweren das Schaffen entsprechender Anreize für Datengebende. Aufgrund von Fragen des Vertrauens und dem Genügen von Neutralitätsanforderungen zeichnet sich bei einer Mehrheit der Pilotprojekte eine Organisationsform, bei der öffentliche Einrichtungen als Träger eine Rolle spielen. Die Gefahr einer Monopolbildung mittels der erfolgreichen Etablierung von DT wurde mehrheitlich nicht gesehen. Im weiteren Studienverlauf werden mögliche Zahlungsmodalitäten und Formen der Bepreisung von Daten sowie die Gefahr kartellrechtlich relevanter Absprachen aufgrund des Datenteilens erarbeitet.

Zu Fragen der **Akzeptanz, Skalierung und des Transfers** von DTM lässt sich feststellen, dass die Akzeptanz bei Datennutzenden ausgeprägter ist als bei Datengebenden. Skalierung und Transfer hängen maßgeblich von der Entwicklung von Standards und Zertifizierungsmechanismen ab. Im weiteren Studienverlauf soll die Frage der Rolle staatlicher Förderung und Infrastruktur in diesem Zusammenhang vertieft untersucht werden. Außerdem wird ein Fokus auf dem Schaffen von Anreizen für Datengebende liege. Bei der Skalierung stellt sich unter anderem die Frage, wie branchenübergreifendes Wachstum und der Zusammenschluss mehrerer DT ermöglicht werden kann. Auch die Anschlussfähigkeit der in den Pilotprojekten erprobten DTM an europäische Initiativen wie Gaia-X wird zu untersuchen sein.